A mutation in T7 RNA polymerase enhances the yield of 5'-guanosine-analog-initiated RNAs

Undergraduates Honors Research Thesis

Presented in Partial Fulfillment of the Requirements for graduation “with Honors Research Distinction in Biochemistry” in the College of Arts and Sciences of The Ohio State University

By

Seth Lyon

The Ohio State University

2018

Project Advisor: Dr. Venkat Gopalan, Department of Chemistry and Biochemistry

Examination committee:

Dr. Venkat Gopalan, Department of Chemistry and Biochemistry

Dr. Kotaro Nakanishi, Department of Chemistry and Biochemistry

Dr. Juan Alfonzo, Department of Microbiology

Abstract

Spectroscopic methods, which are used to establish RNA structure-function relationships, require strategies for post-synthetic, site-specific incorporation of chemical probes into target RNAs. For RNAs larger than 50 nucleotides (nt), the enzymatic incorporation of a nucleoside or nucleoside monophosphate guanosine analog (G-analog) at their 5'-end is routinely achieved by T7 RNA polymerase (T7RNAP)-mediated *in vitro* transcription (IVT) of the appropriate DNA template containing a GTP-initiating class III Φ6.5 promoter. However, when high G-analog:GTP ratios are used to bias G-analog incorporation at the 5'-end, RNA yield is compromised. Here, we show that the use of a T7RNAP Pro266Leu mutant in an IVT with 10:1 thienoguanosine (thG, a fluorescent guanosine surrogate):GTP increased the percent incorporation and yield of 5'-thG-initiated precursor tRNA for a net three-fold gain compared to an IVT with wild-type T7RNAP. We also demonstrate that a one-pot multi-enzyme (OPME) approach, consisting of transcription by T7RNAP Pro266Leu and post-transcriptional cleanup by a polyphosphatase and an exonuclease, led to essentially near-homogeneous 5'-thG-modified transcripts. We also validated the broader utility of our one-pot multi-enzyme approach to initiate RNAs with an azide-bearing G-analog, an undertaking that necessitated the development an RNase P (a tRNA 5′-maturation endonuclease)-based assay to accurately determine the percentage of 5'-azide-initiated RNA. The OPME approach, coupled with use of T7RNAP Pro266Leu, should be of broad utility in generating near homogenous 5'-G-analog RNAs.

Dedication

To my mother, sister, and Richa for always believing in me

and for their relentless support.

Acknowledgments

Foremost, I am forever grateful for the mentorship Dr. Venkat Gopalan has given me. Dr. Gopalan has been instrumental in showing me the wonders of science and in teaching me how to effectively present my research and improve my writing abilities. I cannot thank him enough for all the time and energy he has invested in helping me grow as a scientist. I also want to express my gratitude to the members of the Gopalan laboratory: Dr. Lien B. Lai, Stella Lai, Ila Marathe, Dr. Tien-Hao Chen, Dr. Anindita Sengupta, and Dr. Pradip Biswas for their suggestions, for helping me troubleshoot my experiments, and for their constant willingness to critique my talks when preparing for a poster session.

I thank my mother and sister for all that they have done for me throughout my entire life. They have molded me into the person that I am today, and I am so thankful for their never-ending love and support. I also thank my girlfriend, Richa Agrawal, for being the rock in my life throughout my undergraduate years.

I am also grateful to Drs. Yuhong Zuo and Thomas Steitz (Yale University) for providing the T7RNAP P266L expression vector as well as a detailed purification protocol, and to Dr. Yitzhak Tor (University of California, San Diego) for generously providing thG used in this study and for his valuable feedback. I deeply appreciate the insightful comments and suggestions of Drs. Mikhail Kashlev (NCI), Craig Martin (UMass) and Irina Artsimovitch (OSU). I also thank Prof. E. J. Behrman (OSU) for his encouragement and support. I am also thankful to The Ohio State College of Arts and Sciences for research scholarships, NSF REU grant DBI-1062144, and the Pelotonia Undergraduate Fellowship Program for supporting my work.

Vita

January 27, 1995..... Born – Norfolk, Virginia

2013..... River Valley High School, Marion, OH

2013 to present..... B.S., Biochemistry

Minor, Computer and Information Science

The Ohio State University

Publications

Lyon S, Gopalan V. (2018) A T7 RNA polymerase mutant enhances the yield of 5'-thienoguanosine-initiated RNAs. *ChemBioChem*, **19**: 142-146.

Yue, T., Park, K. H., Reese, B. E., Zhu, H., Lyon, S., Ma, J., and Zhang, M. (2016) Quantifying drug-induced nanomechanics and mechanical effects to single cardiomyocytes for optimal drug administration to minimize cardiotoxicity. *Langmuir*, **32**: 1909-1919.

Fields of Study

Major field: Biochemistry

Minor field: Computer and Information Science

Table of Contents

List of Tables

List of Figures

Abbreviations

α	alpha
λ	wavelength (nanometers)
Φ	phi
°C	degree Celsius
µCi	microcurie
µg	microgram
µL	microliter
ATP	adenosine triphosphate
^{az}G	5'-deoxy-5'-azidoguanosine
CIP	calf intestinal phosphatase
CTP	cytidine triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EPR	electron paramagnetic resonance
G-analog	guanosine analog
GMP	guanosine monophosphate
GTP	guanosine triphosphate
h	hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	isopropyl-ß-D-thiogalactoside
IVT	*in vitro* transcription
LB	Lysogeny Broth
min	minute
mM	millimolar
NaCl	sodium chloride
NAD^+	nicotinamide adenine dinucleotide
nt	nucleotide

OPME	one-pot multi-enzyme
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
PNA	peptide nucleic acid
pre-tRNA	precursor transfer RNA
RNA	ribonucleic acid
rNTP	ribonucleotide triphosphate
rpm	revolutions per minute
RppH	RNA 5'-polyphosphatase
s	second
T4 PNK	T4 polynucleotide kinase
T7RNAP	T7 RNA polymerase
TB	Terrific Broth
thG	thienoguanosine
Tris	tris(hydroxymethyl)aminomethane
UTP	uridine triphosphate
VCE	Vaccina capping enzyme
WT	wild-type
Xrn-1	Exoribonuclease 1

Chapter 1: Introduction*

1.1: The biological importance of RNA

The central dogma of biology, proposed nearly 60 years ago, describes the flow of genetic information from deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) to proteins. Although this framework envisioned RNA as a docile intermediary in information transfer, this notion has been upended by the diversity of RNA functions (e.g., catalysis, chromatin remodeling, structural scaffolds) recognized over the last 30 years.[1,2] These discoveries have motivated a wide gamut of studies spanning RNA's central role in evolution ('RNA world hypothesis')[3,4] to its contribution to diseases.[5] Along with an appreciation of RNA's broad functional scope, there has been progress in developing chemically-modified RNAs that could serve as either drug carriers as well as the design of small-molecule inhibitors to target RNAs implicated in disease.[6-8] Clearly, progress in these efforts to understand and manipulate RNA function requires knowledge of its structure, which in turn necessitates the development of toolkits that will allow facile probing of structure-function relationships of RNAs.[9-13] These toolkits typically seek to generate an RNA with a site-specific biophysical probe that in turn permits investigation of RNA structure under various conditions.

1.2: Methods to incorporate chemical probes into RNA

A suite of post-synthetic strategies is available to introduce chemical modifications to RNAs. Hybridization of a chemically-modified DNA or peptide nucleic acid (PNA) oligomer to an RNA is one approach to indirectly introduce fluorophores or chemical

*: Sections of this chapter are reproduced from Lyon and Gopalan (2018). These sections have benefited from the contributions of both authors.

modifications to an RNA of interest.[9-12,14] While this method is facile and an array of chemically-modified DNA/PNA oligomers are commercially available, it limits the probe to be present at only single-stranded regions of an RNA[9-12,14] and it is possible that binding of the probe interferes with RNA function.

Periodate oxidation offers a route to post-synthetically introduce direct chemical modifications to the 3' terminus of a target RNA.[9,10] This method entails the use of sodium periodate to oxidize the 2',3'-diol at the 3' terminus of the RNA to yield a reactive dialdehyde, which can be subsequently coupled to an amine-bearing fluorophore or other reactive group. Enzymes such as T4 polynucleotide kinase (PNK) and poly(A)-polymerase have also been utilized to introduce chemical modifications at the 5' and 3' termini, respectively, of RNAs. T4 PNK has specifically been used to transfer the γ-phosphorothioate of ATPγS to the 5'-hydroxyl of the target RNA while poly(A)-polymerase adds a 3'-deoxy ATP analog to the 3' hydroxyl of the target RNA.[9,13] Although these methods are typically high yielding, they necessitate an additional processing step.

To date, there have been two reports that demonstrate the use of a capping enzyme from the Vaccina virus, which catalyzes attachment of a m^7G cap the 5'-end of an RNA through a 5'-5' triphosphate bridge, to post-synthetically introduce chemical modifications to the 5'-terminus in a target RNA. The first study utilized the Vaccina Capping Enzyme (VCE) to attach a 3'-anthraniloyl m^7GTP to produce fluorescent RNA and moreover, reported that the capping efficiency was, in some instances, indistinguishable from that of GTP, to generate a fluorescent RNA.[15] The second report used VCE to attach a 3'-biotin-GTP to the 5'-end of target mRNAs and while they did

not report the capping efficiency, they did note that VCE does not accept GMP or 2'-biotin-modified GTP substrates.[16]

The incorporation of a chemical modification during synthesis of an RNA offers a more convenient route to modify RNA. For small RNAs (less than 50 nt), solid-phase synthesis is typically used to site-specifically introduce chemical modifications to both the 5' and 3' termini as well as at an internal position in a target RNA.[9,10] Although solid-phase synthesis allows for direct chemical modification of RNA, it also entails the non-trivial synthesis of a phosphoramidite-bearing NTP analog and is inefficient when generating RNAs larger than 50 nt, a size barrier that is slowly being breached.[17] While the site-specific incorporation of a nucleotide analog during the enzymatic synthesis of RNA through *in vitro* transcription (IVT) of a DNA template containing unnatural DNA base-pairs has been achieved,[18-21] the yield of RNA and/or incorporation of the nucleotide analog is typically poor.[18,20] However, the Ellington laboratory recently reported that use of a T7 RNA polymerase (T7RNAP) with eight mutations, dubbed 'VRS-M5', improved the yield and percent incorporation of a Cy3-bearing ATP analog when transcribing a DNA template with an unnatural base-pair.[21]

1.3: Statement of the problem and workplan

The enzymatic method for site-specifically introducing chemically-modified nucleoside/nucleotide monophosphate analogs at the 5'-terminus of an RNA generated by IVT is particularly attractive due to its low cost and applicability to RNAs of any length. Bacteriophage-derived RNA polymerases (e.g., SP6, T3, T7) have long been

used to synthesize RNAs, with T7RNAP being a well-established and preferred choice.[22-24] A desired RNA is generated by placing the sequence of interest immediately downstream of the 17-bp T7RNAP promoter and by performing an IVT with recombinant, single-polypeptide T7RNAP. This approach, coupled with the Moore-Sharp ligation method, can also be used to generate RNAs with an internal label provided the nucleoside analog has a 5'-hydroxyl available.[25,26] In this scheme, an IVT is used to generate two partial transcripts, which upon ligation will yield the desired full-length RNA. The first partial RNA transcript is designed to span from the 5'-terminus of the full-length RNA to one nucleotide before the desired site of internal modification; the second partial transcript is designed to initiate with the desired modification and terminate at the 3'-end of the full-length RNA. A DNA oligonucleotide, which serves as a splint that is complementary to the junction spanning the two partial transcripts, enables T4 DNA ligase to ligate the two RNAs and yield the full-length RNA with a site-specific internal modification.[25.26]

While there are many T7RNAP promoter variants, the commonly employed class III Φ6.5 promoter allows high transcriptional yields when the RNA is initiated with two guanosines (Gs).[22-24] Several investigators[27-36] have exploited this attribute to bias T7RNAP to initiate with G-analogs (**Table 1.1**), which can support initiation but not elongation since they lack a triphosphate. Despite the remarkable tolerance of T7RNAP to utilize non-native purine mimics as initiators, an intrinsic limitation of this approach is the low yield of RNA resulting from: (i) the use of high G-analog:GTP ratios (up to 100:1) in the IVTs, and (ii) a large fraction of the RNAs initiated with G-analogs finishing up as aborted transcripts. Here, we sought to address these limitations.

In this thesis, I show that the use of a T7RNAP Pro266Leu (P266L) mutant with a decreased propensity for abortive transcription enhances yield and incorporation of the G-analog. Moreover, we used moderate G-analog:GTP ratios (10:1) during IVTs to help enhance the yield of RNAs initiated with GTP and the G-analog, but then employed a post-transcriptional (one-pot) enzymatic clean-up to selectively eliminate 5'-GTP-initiated RNAs. Overall, this simple approach solves a long-standing problem to prepare 5'-modified RNAs with high yield and purity.

I also present a novel RNase P-based assay that was developed to directly assess the percent incorporation of non-fluorescent G-analogs at the 5'-terminus of a pre-tRNA. Importantly, this assay can also be adapted to assess percent incorporation of G-analogs in other long RNAs.

Table 1.1: Previously studied guanosine analogs

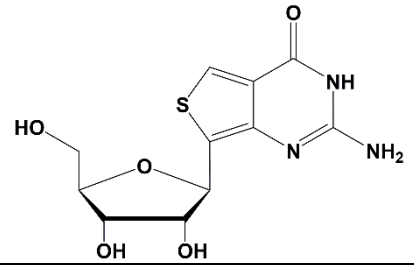
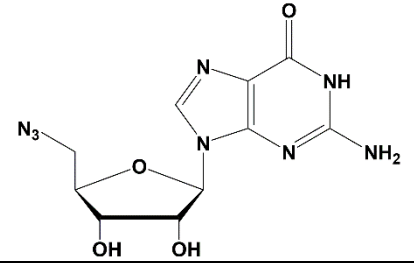
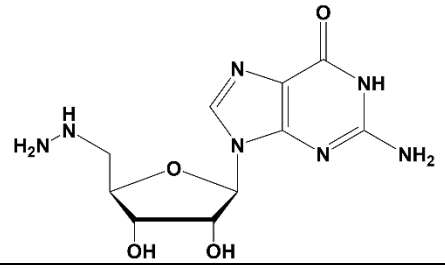
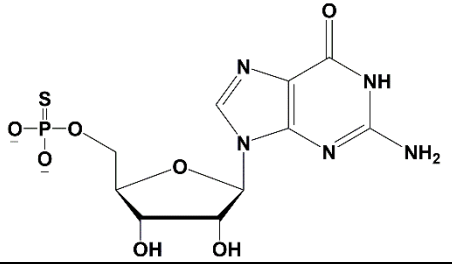
G-analog	G-analog:GTP ratio	Percent incorporation of G-analog	Target RNA length	Reference
i 	5:1	52%	9 nt	27
ii 	4:1	100%	66 nt	28
	5:1	36%	75 nt	29
iii 	31:1	40%	82 nt	30
iv 	8:1	62%	192 nt	31

Table 1.1 continued

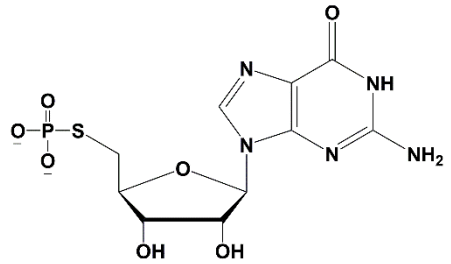
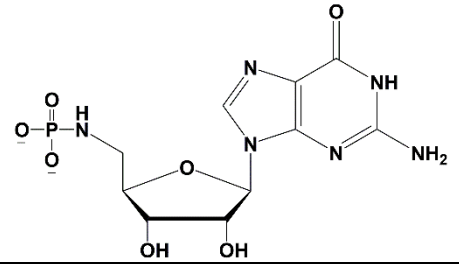
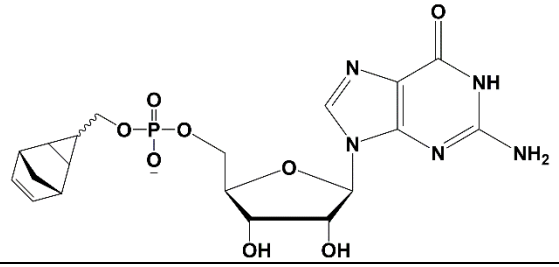
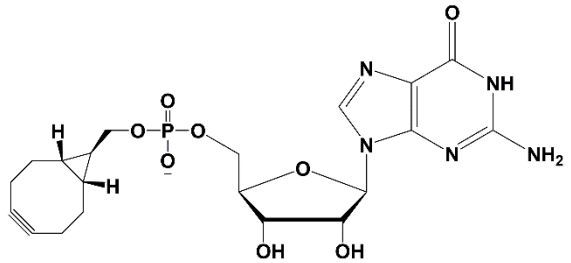
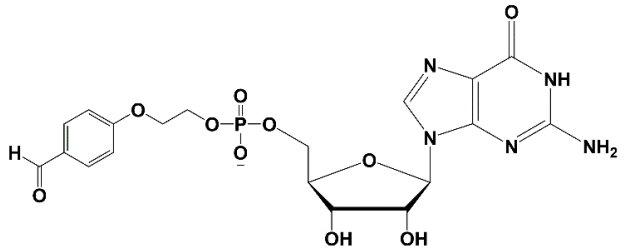
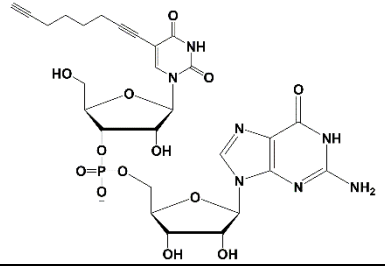
G-analog	G-analog:GTP ratio	Percent incorporation of G-analog	Target RNA length	Reference
v 	8:1	60%	192 nt	32
vi 	18:1	60%	82 nt	33
vii 	10:1	70%	19 nt	34
viii 	10:1	70%	19 nt	34

Table 1.1 continued

G-analog	G-analog:GTP ratio	Percent incorporation of G-analog	Target RNA length	Reference
ix 	10:1	75%	109 nt	35
x 	5:1	80%	25	36

Names of compounds listed above: **i)** thienoguanosine, **ii)** 5'-deoxy-5'-azidoguanosine, **iii)** 5'-deoxy-5'-hydrazinylguanosine, **iv)** 5'-guanosine-5'-monophosphorothioate, **v)** 5'-deoxy-5'-thioguanosine-5'-monophosphorothioate, **vi)** 5'-deoxy-5'-aminoguanosine-*N*-phosphoramidate, **vii-x):** see references for formal chemical names.

1.4: References

[1] Eddy, S. R. (2011) Non-Coding RNA genes and the modern RNA world. *Nat. Rev. Genet*., **2**: 919-929.

[2] Li, L. C. (2014) Chromatin remodeling by the small RNA machinery in mammalian cells. *Epigenetics*, **9**: 45-52.

[3] Joyce, G. F. (2002) The antiquity of RNA-based evolution. *Nature*, **418**: 214-221.

[4] Cech, T. R. (2009) Crawling out of the RNA world. *Cell*, **136**: 599-602.

[5] Cooper, T. A., Wan, L., and Dreyfuss, G. (2009) RNA and Disease. *Cell*, **136**: 777-793.

[6] Saito, Y., Hashimoto, Y., Arai, M., Tarashima, N., Miyazawa, T., Miki, K., Takahashi, M., Furukawa, K., Yamazaki, N., Matsuda, A., Ishida, T. and Minakawa, N. (2014) Chemistry, properties, and *in vitro* and *in vivo* applications of 2'-O-methoxyethyl-4'-thioRNA, a novel hybrid type of chemically modified RNA. *ChemBioChem*, **15**: 1-7.

[7] Shu, Y., Haque, F., Shu, D., Li, W., Zhu, Z., Kotb, M., Lyubchenko, Y., and Guo, P. (2013) Fabrication of 14 different RNA nanoparticles for specific tumor targeting without accumulation in normal organs. *RNA*, **19**: 767-777.

[8] Stelzer, A. C., Frank, A. T., Kratz, J. D., Swanson, M. D., Gonzalez-Hernandez, M. J., Lee, J., Andricioaei, I., Markovitz, D., and Al-Hashimi, H. M. (2011) Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol.*, **7**: 553-559.

[9] Qin, P. Z., and Pyle, A. M. (1999) Site-specific labeling of RNA with fluorophores and other structural probes. *Methods*, **18**: 60-70.

[10] Paredes, E., Evans, M., and Das, S. R. (2011) RNA labeling, conjugation and ligation. *Methods*, **54**: 251-259.

[11] Alexander, S. C., and Devaraj, N. K. (2017) Developing a fluorescent toolbox to shed light on the mysteries of RNA. *Biochemistry*, **56**: 5185-5193.

[12] Rombouts, K., Braeckmans, K., and Remaut, K. (2016) Fluorescent labeling of plasmid DNA and mRNA: gains and losses of current labeling strategies. *Bioconj. Chem.*, **27**: 280-297.

[13] Pagano, J. M., Clingman, C. C., and Ryder, S. P. (2011) Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA*, **17**: 14-20.

[14] Xia, X., Piao, X., and Bong, D. (2014) Bifacial peptide nucleic acid as an allosteric switch for aptamer and ribozyme function. *J. Am. Chem. Soc.*, **136**: 7265-7268.

[15] Gunawardana, D., Domashevskiy, A. V, Gayler, K. R., and Goss, D. J. (2015) Efficient preparation and properties of mRNAs containing a fluorescent cap analog: Anthraniloyl-m^{7}GpppG. *Translation*, **3**: e988538.

[16] Ettwiller, L., Buswell, J., Yigit, E., and Schildkraut, I. (2016) A novel enrichment strategy reveals unprecedented number of novel transcription start sites at single base resolution in a model prokaryote and the gut microbiome. *BMC Genomics*, **17**: 1-14.

[17] Prediger, E. (2017) Long, custom RNA oligos-ultramer RNA oligonucleotides. *DECODED*, www.idtdna.com/pages/education/decoded/article/longcustom-rna-oligos-ultramer-rna-oligonucleotides.

[18] Domnick, C., Eggert, F., and Kath-Schorr, S. (2015) Site-specific enzymatic introduction of a norbornene modified unnatural base into RNA and application in post-transcriptional labeling. *Chem. Commun.,* **51**: 8253-8256.

[19] Lavergne, T., Lamichhane, R., Malyshev, D. A., Li, Z., Li, L., Sperling, E., Williamson, J. R., Millar, D. P., and Romesberg, F. E. (2016) FRET characterization of complex conformational changes in a large 16S ribosomal RNA fragment site-specifically labeled using unnatural base pairs. *ACS Chem. Biol.*, **11**: 1347-1353.

[20] Eggert, F., and Kath-Schorr, S. (2016) A cyclopropene-modified nucleotide for site-specific RNA labeling using genetic alphabet expansion transcription. *Chem. Commun.*, **52**: 7284-7287.

[21] Kimoto, M., Meyer, A. J., Hirao, I., and Ellington, A. D. (2017) Genetic alphabet expansion transcription generating functional RNA molecules containing a five-letter alphabet including modified unnatural and natural base nucleotides by thermostable T7 RNA polymerase variants. *Chem. Commun*., **53**: 12309-12312.

[22] Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.*, **180**: 51-62.

[23] Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**: 8783-8798.

[24] Masquida, B., Beckert, B., and Jossinet, F. (2010) Exploring RNA structure by integrative molecular modelling. *N. Biotechnol*., **27**: 170-183.

[25] Moore, M.J. and Sharp, P.A. (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science*, **256**: 992-997.

[26] Kershaw, C. J., and Keefe, R. T. O. (2012) Splint ligation of RNA with T4 DNA ligase. *Methods Mol. Bio*., **941**: 257-269.

[27] Li, Y., Fin, A., McCoy, L., and Tor, Y. (2017) Polymerase-mediated site-specific inorporation of a synthetic fluorescent isomorphic G surrogate into RNA. *Angew. Chem. Int. Ed.*, **56**: 1303-1307.

[28] Paredes, E. and Das, S. R. (2011) Click chemistry for rapid labeling and ligation of RNA. *ChemBioChem*, **12**: 125-131.

[29] Lee, G. H., Lim, H. K., Jung, W., and Hah, S. S. (2012) Incorporation efficiency of 5'-azido-5'-deoxyguanosine into 5'-terminus of RNA for preparation of azido-functionalized RNA. *Bull. Korean Chem. Soc.*, **33**: 3861-3863.

[30] Skipsey, M., Hack, G., Hooper, T. A., Shankey, M. C., Conway, L. P., Schroder, M., and Hodgson, D. R. W. (2013) 5'-Deoxy-5'-hydrazinylguanosine as an initiator of T7 RNA polymerase-catalyzed transcriptions for the preparation of labeling-ready RNAs. *Nucleos. Nucleot. Nucl.*, **32**: 670-681.

[31] Zhang, L., Sun, L., Cui, Z., Gottlieb, R. L., and Zhang, B. (2001) 5'-Sulfhydryl-modified RNA: initiator synthesis, *in vitro* transcription, and enzymatic incorporation. *Bioconjug. Chem.*, **12**: 939-948.

[32] Zhang, B., Cui, Z., and Sun, L. (2001) Synthesis of 5'-deoxy-5'-thioguanosine-5'-monophosphorothioate and its incorporation into RNA 5'-termini. *Org. Lett.*, **3**: 275-278.

[33] Williamson, D., Cann, M. J., and Hodgson, D. R. W. (2007) Synthesis of 5'-amino-5'-deoxyguanosine-5'-N-phosphoramidate and its enzymatic incorporation at the 5'-termini of RNA molecules. *Chem Commun.*, **47**: 5096-5098.

[34] Ameta, S., Becker, J., and Jäschke, A. (2014) RNA-peptide conjugate synthesis by inverse-electron demand Diels-Alder reaction. *Org.Biomol. Chem.*, **12**: 4701-7.

[35] Samanta, A., Krause, A., and Jäschke, A. (2014) A modified dinucleotide for site-specific RNA-labelling by transcription priming and click chemistry. *Chem. Commun.*, **50**: 1313-1316.

[36] Pfander, S., Fiammengo, R., Kirin, S. I., Metzler-Nolte, N., and Jäschke, A. (2007) Reversible site-specific tagging of enzymatically synthesized RNAs using aldehyde-hydrazine chemistry and protease-cleavable linkers. *Nucleic Acids Res.*, **35**: 1-8.

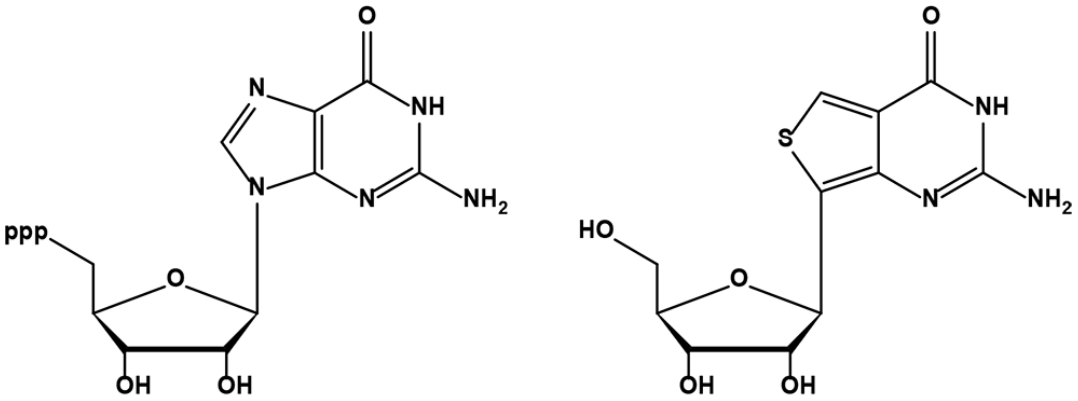
Chapter 2: A T7 RNA polymerase mutant enhances the yield of 5'-thienoguanosine-initiated RNAs**

2.1: Background and rationale

To obtain a high yield of full-length RNAs containing a G-analog at the 5'-end, we chose a test G-analog and examined a set of modified IVT conditions. Three reasons inspired our choice of thienoguanosine (thG; **Figure 2.1**), a fluorescent guanosine surrogate,[1,2] as the pilot analog. First, thG permits facile fluorescence detection[1] post-synthesis, thus allowing us to rapidly examine different IVT conditions and fine-tune variables as needed. Second, optimizing 5'-incorporation is expected to have payoffs for site-specific (internal) modification because 5'-thG-modified RNAs can be 5'-phosphorylated and ligated to another RNA to yield the desired longer RNA with an internal thG.[2] Finally, given the value of emissive nucleoside analogs as *bona fide* isomorphs, methods to enhance incorporation of these isomorphs into RNAs will greatly aid mechanistic and spectroscopic studies, as was demonstrated elegantly in a recent study of the hammerhead ribozyme.[2] As a model RNA for our IVTs, we used a 155-nt precursor-tRNA^{Cys} (pre-tRNA^{Cys}; **Figure 2.2**).

We considered various modifications to the IVT conditions to increase incorporation of the G-analog. Foremost, based on the premise that poor IVT yields arise from the rapid depletion of limited GTP, which is typically used at low concentrations to favor incorporation of the modified G-analog, Katie Adib, a previous undergraduate in the Gopalan laboratory, used a phased-addition strategy wherein GTP was supplemented at timed intervals. Although this approach engendered 5'-incorporation gains, the increase was not uniform for different G-analogs.[3] Second, we

**: Chapter 2 has been reproduced verbatim from Lyon and Gopalan (2018).



Guanosine triphosphate **Thienoguanosine**

Figure 2.1: Structures of guanosine triphosphate and thienoguanosine.

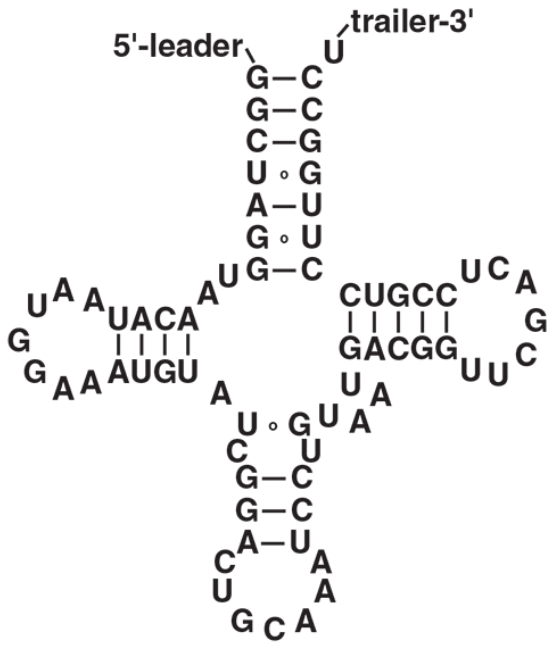


Figure 2.2: Secondary structure of pre-tRNACys, the model RNA used in this study. The sequences of the 5’-leader and 3’-trailer are provided below the secondary structure.

included varying concentrations of Mn^{2+} in the IVT to influence T7RNAP fidelity[4,5] and thereby favor the G-analog over GTP, but these attempts were unsuccessful. Results from these studies revealed that abortive transcription is a critical roadblock, and inspired us to consider the use of T7RNAP mutants that have been shown to increase incorporation of various 2'-modified NTP-analogs into RNAs synthesized during IVTs.[6-9] From this panel of mutants, we chose T7RNAP P266L because kinetic and footprinting studies revealed that this mutant exhibits a delayed transition from transcriptional initiation to elongation and generates fewer aborted transcripts, an attribute that stems from the enhanced stability of the longer DNA:RNA hybrid in its initiation complex.[10-13] We sought to test the postulate that the T7RNAP P266L mutant would decrease aborted transcripts typically observed when 5'-G-analogs are used with T7RNAP wild-type (WT; **Figure 2.3**). The problem of low G-analog incorporation, however, necessitates a combination of transcriptional and post-transcriptional strategies since a fraction of transcripts will always be initiated with GTP. Therefore, we investigated the utility of a one-pot multi-enzyme (OPME) approach wherein a clean-up step (post-transcription with T7RNAP P266L) with a polyphosphatase and an exonuclease is used in tandem to selectively degrade the 5'-GTP-initiated RNAs and yield near-homogeneous 5'-thG-initiated RNAs (**Figure 2.4**).

2.2: Materials and methods

2.2.1: Overexpression and purification of T7RNAP WT

The purification of T7RNAP WT exploited immobilized metal affinity-chromatography and was performed essentially as described earlier,[14] albeit with minor changes as outlined

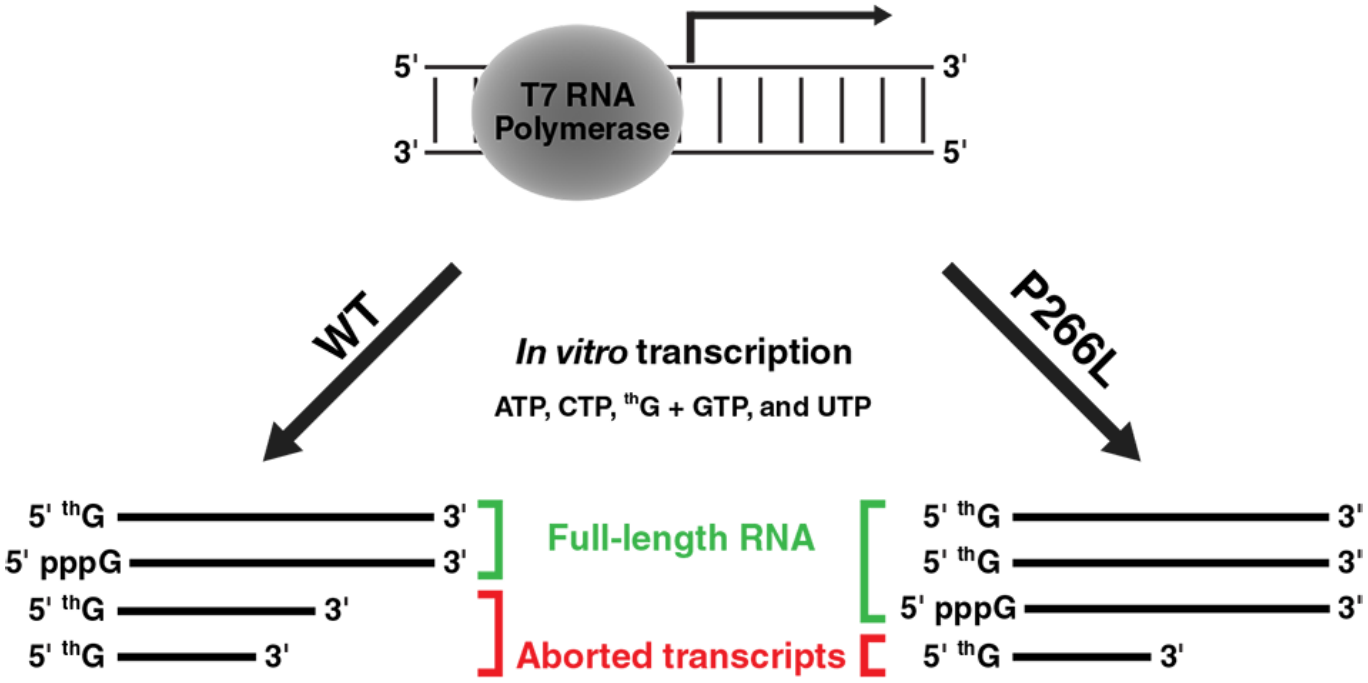


Figure 2.3: Schematic of the expected outcomes from IVTs with a fixed ratio of thG:GTP and performed with either T7RNAP WT or the P266L mutant. This illustration depicts the qualitative gains with respect to abortive transcription (see **Figure 2.8**).

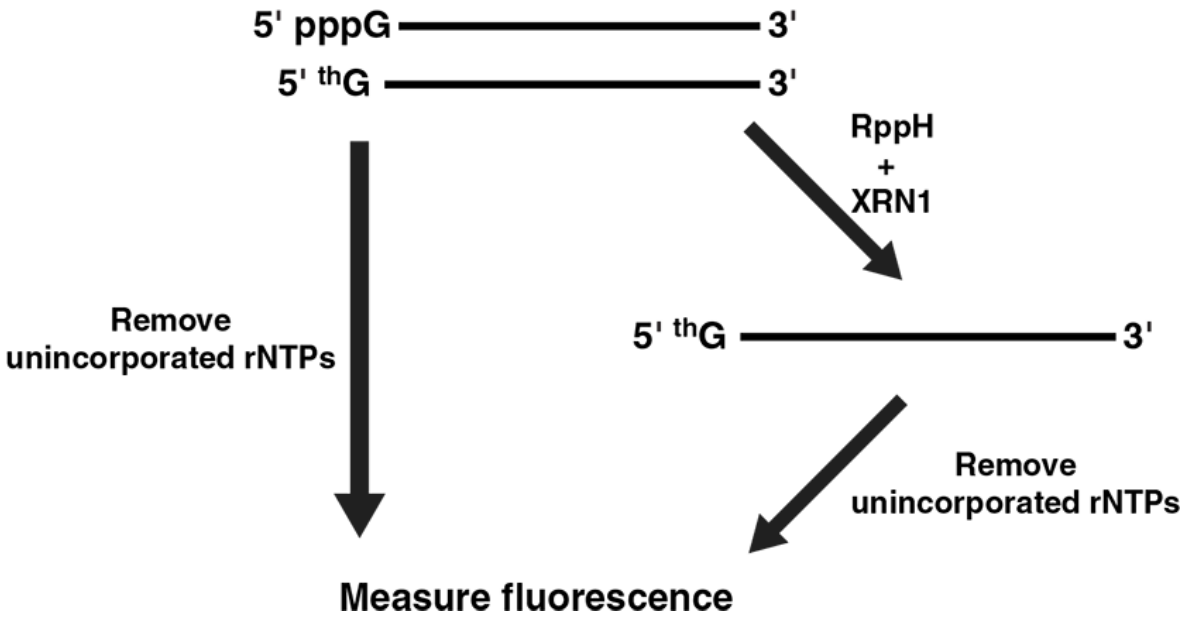


Figure 2.4: Comparison of a typical IVT versus the one-pot multi-enzyme approach designed to eliminate 5'-GTP-initiated RNAs.

below. *Escherichia coli* BL21(DE3) cells were transformed with two plasmids: pQE9T7 (amp^r) and pREP4 (kan^r) which encode $(His)_6$-T7RNAP WT and *lac* repressor, respectively, and grown on a lysogeny broth (LB) agar plate containing 100 µg/mL carbenicillin and 35 µg/mL kanamycin at 37°C for 16 h. A single colony was picked to inoculate 5 mL of LB containing 100 µg/mL carbenicillin and 35 µg/mL kanamycin, and grown at 37°C for 16 h with shaking at 200 rpm. The 5-mL culture was then used to inoculate 500 mL of LB containing 100 µg/mL carbenicillin and 35 µg/mL of kanamycin, and grown at 37°C with shaking at 200 rpm until OD_{600} ~0.6. The culture was induced by addition of isopropyl-ß-D-thiogalactoside (IPTG) to a final concentration 1 mM and then grown at 37°C for an additional 2.5 h. Cells were harvested by centrifugation at 6,000 g for 20 min, and the pellets stored at -80°C until further use.

A 500-mL cell pellet was thawed on ice and then resuspended in buffer A [100 mM sodium phosphate (pH 7.7), 300 mM NaCl, 0.5 mM DTT, 1 mM PMSF, 2 mini EDTA-free protease inhibitor tablets (Roche)]. Cells were lysed by sonication (60% amplitude, 12-min of alternating cycles of 4 s ON and 10 s OFF; Cole Parmer 130-Watt Ultrasonic Processor model GEX130) and the crude lysate was subjected to centrifugation at 20,000 g for 1 h at 4°C. The supernatant obtained was filtered through a 0.45 µm syringe filter. After addition of 20 mM imidazole to the filtered supernatant, it was loaded onto a pre-equilibrated 5-mL HisTrap HP (GE Healthcare) column. An ATKA FPLC (GE Healthcare) was used for the subsequent elution. The column was washed with 50 mL buffer B [100 mM sodium phosphate (pH 7.7), 300 mM NaCl, 0.5 mM DTT, 1 mM PMSF, 20 mM imidazole], and T7RNAP WT was eluted using a 100-ml linear gradient from 20-300 mM imidazole at a flow rate of 3 mL/min. T7RNAP WT

typically eluted between 170-210 mM imidazole. Peak fractions, as judged by SDS-PAGE analysis, were pooled and subsequently dialyzed against 500 mL of buffer C [50 mM sodium phosphate (pH 7.7), 100 mM NaCl, 1 mM DTT, 0.1 mM PMSF, 1 mM EDTA and 50% (v/v) glycerol] at 4°C overnight after one buffer exchange after 2 h. Aliquots of the purified T7RNAP (dialysate) were stored at -80°C until further use in in vitro transcriptions (IVTs).

2.2.2: Overexpression and purification of T7RNAP P266L

The overexpression and purification of $(His)_6$-T7RNAP P266L was based on a protocol kindly provided by the laboratory of Prof. Thomas A. Steitz, Yale University,[10] with some notable modifications (exclusion of a polishing step as well as dialysis and final preparative steps). *E. coli* BL21(DE3) cells were transformed with pBH161 (a gift from the Steitz laboratory), which encodes $(His)_6$-T7RNAP P266L, and grown on an LB-agar plate containing 50 µg/mL carbenicillin at 37°C for 16 h. A single colony was picked to inoculate 5 mL of Terrific Broth (TB) containing 50 µg/mL carbenicillin and the culture was grown at 37°C for 16 h with shaking at 200 rpm. This 5-mL seed culture was added to 500 mL of TB containing 50 µg/mL carbenicillin and grown at 37°C with shaking at 200 rpm until OD_{600} ~1. Next, carbenicillin was supplemented to a final concentration of 100 µg/mL and the culture induced with 1 mM IPTG at 37°C for 4 h. Cells were harvested (as described above for WT) and the pellets stored at -80°C.

For purification of T7RNAP P266L, a 500-mL cell pellet was thawed on ice and then resuspended in buffer D [10 mM Tris-HCl (pH 8.0), 375 mM NaCl, 0.5 mM DTT, 1

mM PMSF, 2 mini EDTA-free protease inhibitor tablets (Roche)]. Subsequent steps mirrored the T7RNAP WT procedure described above with the exception of use of a 1-mL HisTrap (GE Healthcare) column. After washing with 10 mL of buffer E [10 mM Tris-HCl (pH 8.0), 375 mM NaCl, 0.5 mM DTT, 1 mM PMSF, 20 mM imidazole], T7RNAP P266L was eluted using a 12-ml linear gradient from 20-200 mM imidazole. T7RNAP P266L typically eluted between 140-180 mM imidazole. Peak fractions, as assessed by SDS-PAGE analysis, were pooled and subsequently dialyzed in 500 mL of buffer F [100 mM sodium phosphate (pH 7.7), 200 mM NaCl, 2 mM DTT, 0.2 mM PMSF, 2 mM EDTA] at 4°C for 3 h with buffer exchanges every hour. An equal volume of glycerol was added to the dialysate and the solution was mixed at 4°C using a nutator. Aliquots of the purified T7RNAP P266L [in 50% (v/v) glycerol)] were stored at -80°C until further use.

2.2.3: In vitro transcription reactions

A PCR DNA template containing a class III Φ6.5 promoter and gene encoding *Arabidopsis thaliana* mitochondrial precursor $tRNA^{Cys}$ was prepared as previously described.[15] This pre-tRNA has a 55-nt 5'-leader and a 23-nt 3'-trailer (see **Figure 2.2**), and was chosen due to its availability and use in our laboratory for studies of RNase P. One hundred-µl IVTs were performed at 37°C for 5 h in 80 mM HEPES (pH 7.5), 1 mM spermidine, 33 mM $MgCl_2$, 5% (v/v) DMSO, 3 mM ATP, 3 mM CTP, 3 mM UTP, 0.2 U of thermostable inorganic pyrophosphatase [New England Biolabs (NEB), Ipswich, MA], 400 ng of DNA template and approximately 2 µg of either T7RNAP WT or P266L. IVTs containing ^{th}G were performed using 4.8 mM ^{th}G and 0.48 mM GTP (10:1 ratio). Post-

transcription, the DNA template was degraded by addition of 10 U of DNase I (Roche, Basel, Switzerland) and incubation at 37°C for 30 min.

2.2.3: Post-transcriptional removal of 5'-ppp RNAs

For the OPME approach, 5 U of the RppH polyphosphatase (NEB) and NaCl (to a final concentration of 50 mM) was added to the DNase I-treated IVT reaction and incubated for 1.5 h at 37°C. Subsequently, 1 U of the Xrn-1 exonuclease (NEB) and NaCl (to a final concentration of 100 mM) was added to the reaction and incubated for 1.5 h at 37°C. To ensure higher Xrn-1 activity, we increased the temperature to 45°C for 5 s every 5 min; these reactions were performed using a thermal cycler.

Post-transcription and RNA processing, all reactions were subjected to a phenol-chloroform extraction and subsequently to a Zymo Clean and Concentrator^{TM}-25 (Zymo Research, Irvine, CA) to remove unincorporated nucleotides and short, abortive transcripts <18 nt in length (**Figure 2.5**). Total RNA yield was determined using Abs_{260} measurements on a NanoDrop 2000c spectrophotometer and the extinction coefficient of pre-tRNA^{Cys}.

2.2.4: Fluorescence measurements to determine percent incorporation of ^{th}G

To generate a standard curve, increasing concentrations of ^{th}G in 20 mM Tris-HCl (pH 7.5) were aliquoted into a 384-well black microplate (Nunc) and fluorescence intensity was recorded using a Tecan M1000 microplate reader (λ_{exc}, 380 nm; λ_{em}, 452 nm).

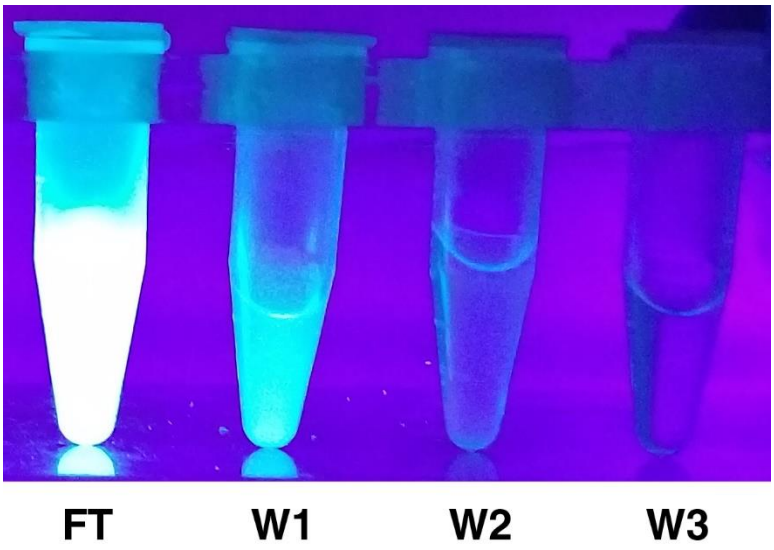


Figure 2.5: Efficacy of the Zymo Clean and Concentrator^{TM}-25 approach to remove unincorporated thG. Tubes were illuminated at 312 nm using a Fotodyne UV transilluminator to visualize the removal of free thG. FT: flow-through; W1: wash using the RNA prep buffer (400 µl); and W2: wash using the RNA wash buffer (700 µl). W3: wash using the RNA wash buffer (400 µl). Short RNAs <18 nt are expected to be removed by this desalting approach using the recommended volumes (according to the manufacturer).

Relative fluorescence units (RFU) were determined by subtracting the background fluorescence of blanks lacking thG, and RFU was then plotted against the concentration of thG to generate a linear curve-fit (Excel). This standard curve was subsequently used to determine the concentration of thG in pre-tRNA^{Cys} (**Figure 2.6**). To investigate if there was any quenching associated with thG present at the 5' terminus of pre-tRNA^{Cys}, we compared the fluorescence of thG-initiated pre-tRNA^{Cys} before and after alkaline hydrolysis and found that the fluorescence values were nearly indistinguishable (data not shown). This finding lends confidence to the estimates of % 5'-modified RNAs that were determined from the thG-based standard curve.

2.3: Results

Prior to comparisons of T7RNAP WT and P2666L mutant, we first confirmed the use of roughly equivalent amounts of these variants in our IVTs (**Figure 2.7**). Because of our interest in 5'-incorporation of 5'-deoxy-5'-azido-guanosine (^{az}G), we had performed IVTs with this analog and obtained data indicating that percent incorporation was higher when the G-analog:GTP ratio was 10:1 compared to 4:1, regardless of whether we used the WT or the P266L mutant. This conclusion is also consistent with a previous report that used ^{az}G.[16] Additionally, by using an IVT spiked with an α-[^{32}P]-labeled rNTP, we found that the fraction of short aborted transcripts relative to full-length was significantly lower with T7RNAP P266L compared to the WT when ^{az}G was the analog (see **Chapter 3**); we confirmed this trend using an IVT performed with thG as well (**Figure 2.8**).

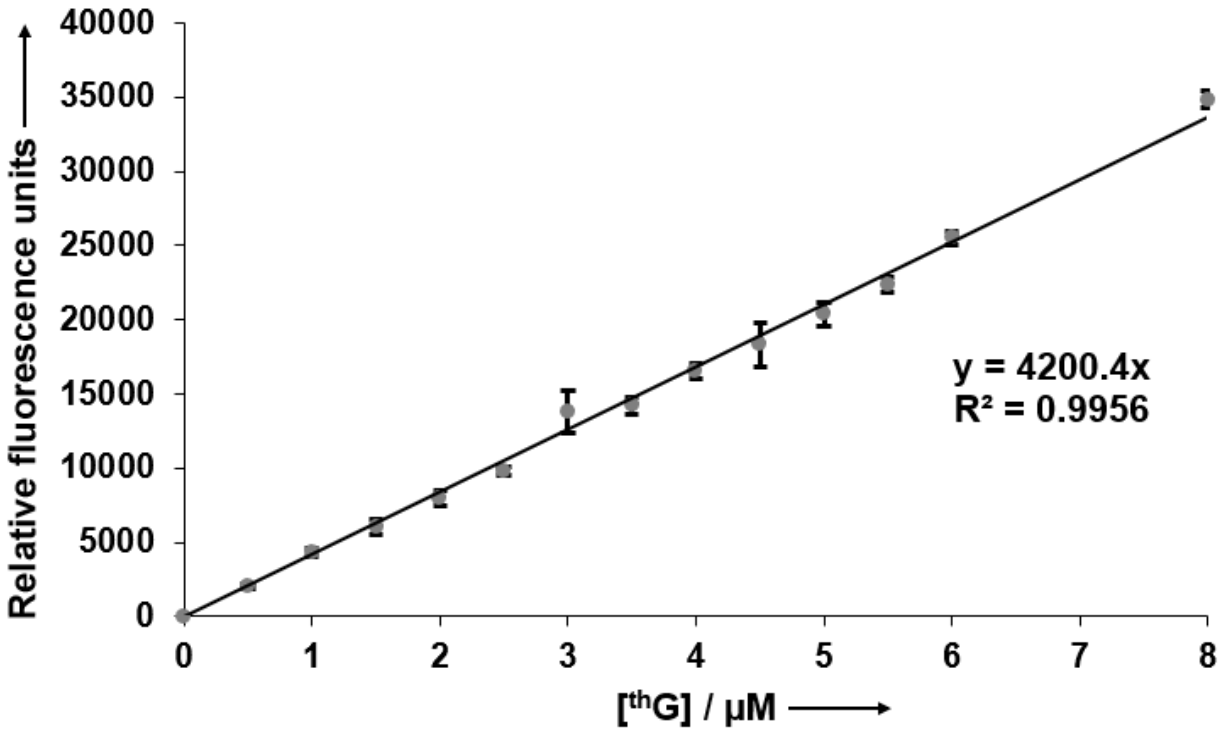


Figure 2.6: Standard curve used to determine percent incorporation of thG into pre-tRNA^{Cys}. Mean and standard deviation values were calculated from three independent measurements.

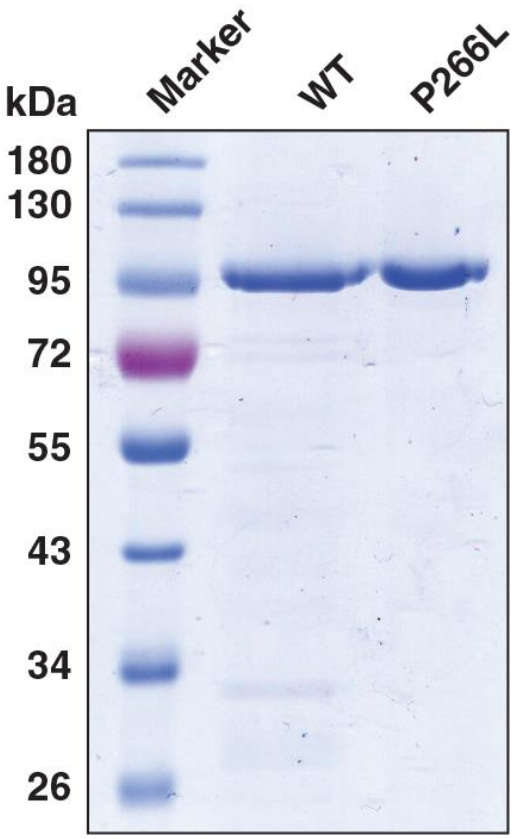


Figure 2.7: Purity of recombinant T7RNAP used in this study, as demonstrated by SDS-PAGE (10%, w/v, polyacrylamide) analysis.

Figure 2.8: T7 RNAP P266L yields fewer abortive transcripts than WT when IVTs are performed with thG. IVTs to generate pre-tRNA^{Cys} were performed using either T7 RNAP WT or P266L as described in the main text, except each reaction also contained 10 µCi of α-^{32}P-GTP (Perkin Elmer, Shelton, CT). Transcription was terminated by addition of a loading dye containing 7 M urea and 20% (v/v) phenol. One-tenth of the 20 µL IVT reaction was loaded on a 7% (w/v) polyacrylamide/7 M urea gel (40 cm length x 20 cm width), and electrophoresed for 75 min at 30 mA. The gel was then exposed for 2 h to a phosphorimager screen, and IVT products were visualized using an Amersham Typhoon 5 Biomolecular Imager (GE Healthcare). Based on quantitation performed using ImageQuant 5.1 software, there is a 1.3-fold increase in the full-length pre-tRNA^{Cys} and a corresponding decrease in aborted transcripts.

Using a 10:1 ratio of thG:GTP, we determined that T7RNAP WT initiates 40% of pre-tRNA^{Cys} transcripts with thG affording a total RNA yield of 37 µg from a 100 µl IVT (**Figures 2.9, 2.10, and 2.11**). However, when T7RNAP P266L was employed in the IVT, thG percent incorporation increased to 85% with a total RNA yield of 50 µg (**Figures 2.9, 2.10, and 2.11**). Thus, the net gain of 5'-thG-modified RNA is ~3-fold. While visual inspection post-agarose gel electrophoresis revealed that the IVT yield is higher with the P266L mutant compared to the WT (**Figure 2.9**), the percent incorporation of thG in pre-tRNA^{Cys} was accurately determined by first eliminating the unincorporated thG and then measuring the fluorescence in the final RNA product. The efficacy of the desalting column in removing the unincorporated thG was evident from examining the washes on a trans-illuminator (λ_{exc}, 312 nm); the ability to easily monitor thG is an appealing feature in its use and allows for rapid visual comparison of its incorporation efficiency into RNA across different IVT conditions (**Figures 2.5 and 2.10**). While use of the desalting column was essential for complete removal of unincorporated thG, we consistently recovered only ~70% of the RNA input to the column. Replacing this step with an alternative that results in a lower loss will be desirable.

To synthesize 5'-NAD-capped RNA, Jäschke and coworkers[17] employed an elegant multi-pot, chemo-enzymatic approach. Post-IVT, the RNA product was subjected to a polyphosphatase treatment to remove the pyrophosphate and yield 5'-GMP transcripts, which were then reacted with nicotinamide mononucleotide-phosphorimidazolide. Any unreacted 5'-GMP transcripts were then removed by Xrn-1. We have now extended this approach to enhance the fraction of 5'-thG-pre-tRNA^{Cys}

albeit with an OPME set-up. Post-IVT, RppH was utilized to convert 5'-GTP-pre-tRNA^{Cys} into 5'-GMP-pre-tRNA^{Cys} that was then degraded using commercially available Xrn-1. This exonuclease was not expected to act on 5'-thG-pre-tRNA^{Cys} given the absence of a 5'-monophosphate. An advantage of the OPME approach is that these enzymes are added post-IVT with no intervening processing steps. Our OPME method resulted in 40 µg of near-homogenous (~95%) 5'-thG-pre-tRNA^{Cys} from a 100-µl IVT (**Figure 2.11**). That a small fraction of the 5'-GTP initiated pre-tRNA^{Cys} was not removed indicates that the activities of RppH and Xrn-1 may not be optimal under our OPME reaction conditions, but these can be tweaked through customized optimization trials.

Owing to the poor solubility of guanosine in water, the stock of thG was prepared in DMSO. The inclusion of 20% DMSO in IVTs was reported to afford an increase in RNA yield and fewer variations in 3'-termini (lower 3'-heterogeneity).[18] To test the idea that inclusion of DMSO in the IVT might similarly influence the bias of T7RNAP to initiate RNAs with thG, we performed IVTs of pre-tRNA^{Cys} as described above except in the presence or absence of 5% DMSO. However, we did not observe any significant effect on RNA yield or incorporation of thG upon inclusion of DMSO (**Figure 2.12**).

2.4: Discussion

Our results show that the T7RNAP P266L mutant enhances the total RNA yield and percent of 5'-thG-initiated RNAs, with additional gains resulting from the post-IVT use of RppH and Xrn-1 to selectively remove the 5'-GTP-initiated RNAs. The findings with thG reaffirm the increased tolerance of T7RNAP P266L mutant to initiate RNA transcripts

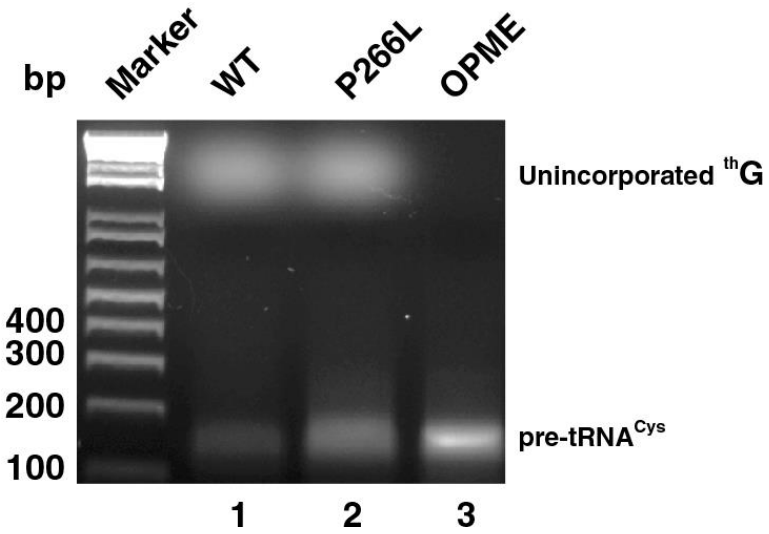


Figure 2.9: One aliquot (1 µL) from 100 µL IVT, performed with either T7RNAP WT (lane 1) or P266L (lane 2), was electrophoresed on a 2% (w/v) agarose gel and stained with ethidium bromide; samples were examined post-DNase I treatment. Lane 3 contains 1 µg of the final pre-tRNA^{Cys} that was generated by using the OPME (post-IVT RppH and Xrn-1) approach. Unincorporated thG does not electrophorese far into the agarose gel because it lacks a negative charge.

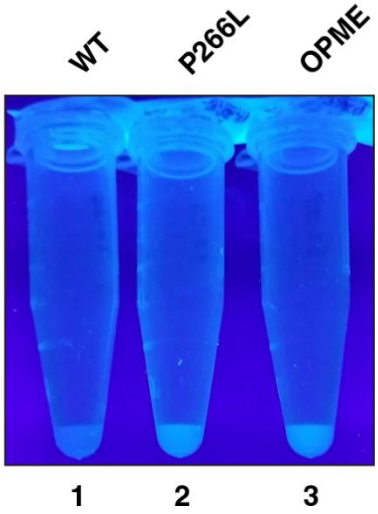


Figure 2.10: Fluorescence (λ_{ex} =312 nm) of 5 mm RNA obtained after IVT with T7RNAP WT (1) or P266L (2) or after using the OPME method with P266L (3).

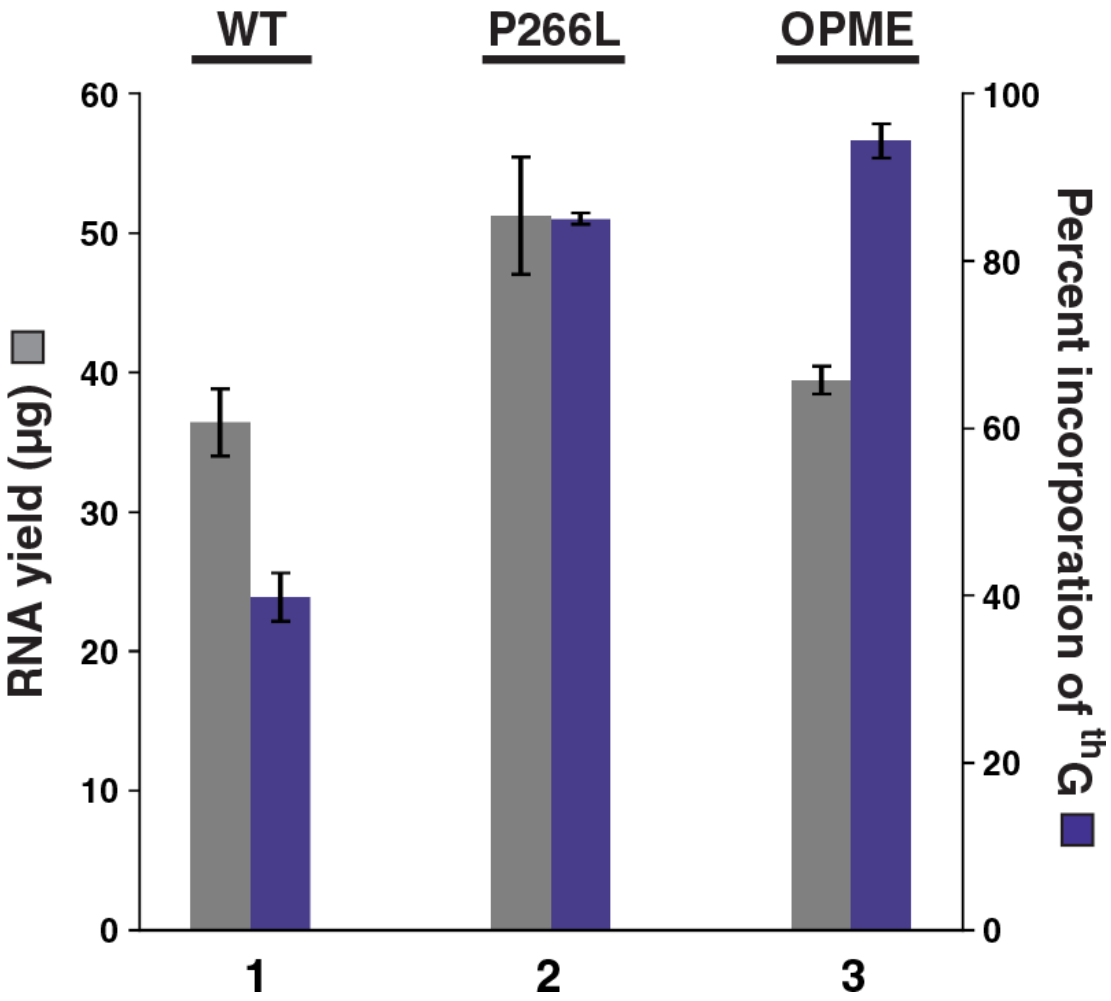


Figure 2.11: RNA yields (gray) and percentage incorporation of thG (purple) with T7RNAP WT (1) or P266L (2) or using the OPME method with P266L (3). Mean and standard deviation values were calculated from three independent measurements.

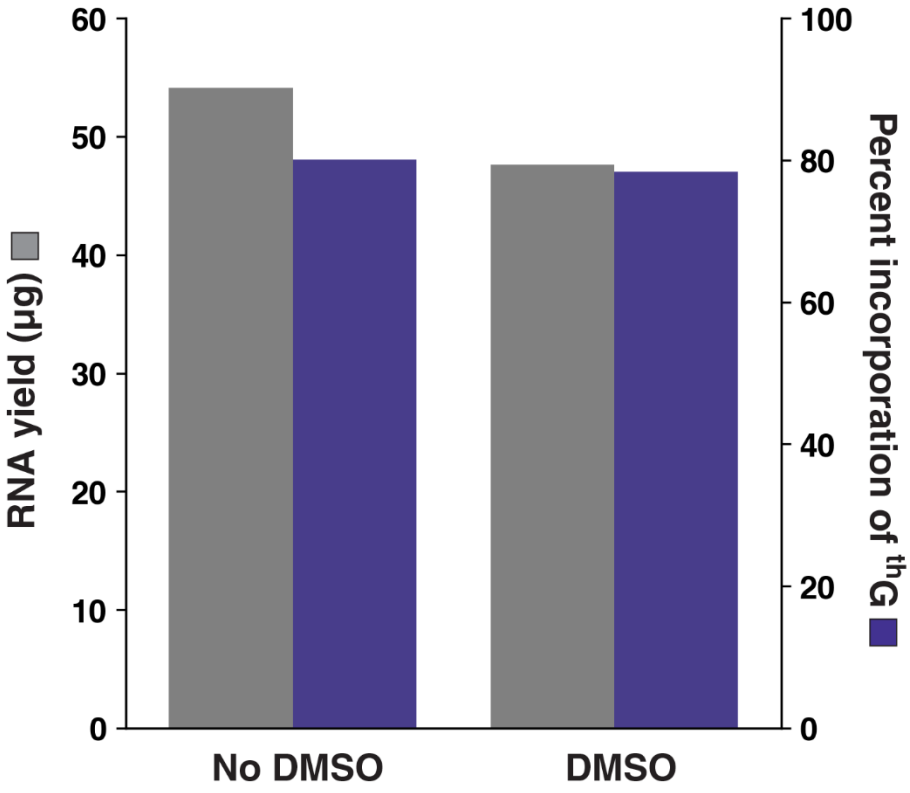


Figure 2.12: RNA yield and percent incorporation of thG into pre-tRNA^{Cys} using T7RNAP P266L in the absence or presence of 5% (v/v) DMSO. The data reported here are from a single IVT trial.

with modified nucleosides, an attribute noted before while investigating use of modified 2'-NTPs.[8] Given the value of thG as a probe of RNA folding/dynamics and RNA function,[2] the facile synthesis of 5'-thG-initiated RNAs should increase the use of this guanosine isomorph. Moreover, we expect that the OPME approach described here will improve percent incorporation and yields of other G-analogs at the 5'-end of RNAs of any length and sequence.

Because the total RNA yield (50 µg/100 µl IVT) of pre-tRNA^{Cys} from IVTs containing thG was ~2.5-fold lower than IVTs containing only 3 mM rNTPs, we sought two different approaches to bridge this gap. First, we explored the use of a mutant T7RNAP containing 10 mutations (termed 'RGVG-M6'), including P266L, which exhibited a 25-fold increase in RNA yield of fully 2'-modified RNA over other commonly used mutant T7RNAPs.[9] When we purified this mutant and tested the recombinant version in IVTs, we did not observe an increase in the total yield of 5'-thG-initiated RNAs. Second, we examined the effects of retaining the [thG]:[GTP] ratio but increasing their absolute concentrations on T7RNAP P266L's activity. Indeed, changing [thG]:[GTP] from 4.8 mM:0.48 mM to 10 mM:1 mM led to doubling of the yield of total pre-tRNA^{Cys}, although percent incorporation of thG decreased by one-fourth (duplicate experiments, data not shown). A modest compromise in the incorporation of thG is offset by the yield payoff especially since the OPME approach helps selectively eliminate the GTP-initiated RNA. Due to the template-dependent variability in IVT yields, pilot tests of different rNTP concentrations in IVTs as well as different [thG]:[GTP] ratios will be gainful for any new thG-initiated RNA.

2.5: References

[1] Shin, D., Sinkeldam, R. W., and Tor, Y. (2011) Emissive RNA alphabet. *J. Am. Chem. Soc.*, **133**: 14912-14915.

[2] Li, Y., Fin, A., McCoy, L., and Tor, Y. (2017) Polymerase-mediated site-specific inorporation of a synthetic fluorescent isomorphic G surrogate into RNA. *Angew. Chem. Int. Ed.*, **56**: 1303-1307.

[3] Adib, K. (2015) A phased-addition strategy enhances the yield of RNAs obtained by *in vitro* transcription with modified transcriptional initiators. *Undergraduate honors thesis*, The Ohio State University, Columbus, OH.

[4] Tabor, S. and Richardson, C. C. (1986) Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.*, **86**: 4076-4080.

[5] Walmacq, C., Kireeva, M. L., Irvin, J., Nedialkov, Y., Lubkowska, L., Malagon, F., Strathern, J. N., and Kashlev, M. (2009) Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II*. J. Biol. Chem.*, **284**: 19601-19612.

[6] Kotkowiak, W., Pasternak, A., and Kierzek, R. (2016) Studies on transcriptional incorporation of 5'-N-triphosphates of 5'-amino-5'-deoxyribonucleosides. *PLoS One*, **11**: e0148282.

[7] Padilla, R. and Sousa, R. (2002) A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs. *Nucleic Acids Res.*, **30**: e138.

[8] Sochor, F., Silvers, R., Muller, D., Richter, C., Furtig, B., and Schwalbe, H. (2016) ^{19}F-labeling of the adenine H2-site to study large RNAs by NMR spectroscopy. *J. Biomol. NMR*, **64**: 63-74.

[9] Meyer, A. J., Garry, D. J., Hall, B., Byrom, M. M., McDonald, H. G., Yang, X., Yin, Y. W., and Ellington, A. D. (2015) Transcription yield of fully 2'-modified RNA can be increased by the addition of thermostabilizing mutations to T7 RNA polymerase mutants. *Nucleic Acids Res.*, **43**: 7480-7488.

[10] Durniak, K. J., Bailey, S., and Steitz, T. A. (2008) The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. *Science*, **322**: 553-557.

[11] Guillerez, J., Lopez, P. J., Proux, F., Launay, H., and Dreyfus, M. (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. U.S.A.*, **102**: 5958-5963.

[12] Ramirez-Tapia, L. E. and Martin, C. T. (2012) New insights into the mechanism of initial transcription: the T7 RNA polymerase mutant P266L transitions to elongation at longer RNA lengths than wild type. *J. Biol. Chem.*, **287**: 37352-37361.

[13] Tang, G. Q., Nandakumar, D., Bandwar, R. P., Lee, K. S., Roy, R., Ha, T., and Patel, S. S. (2014) Relaxed rotational and scrunching changes in P266L Mutant of T7 RNA polymerase reduce short abortive RNAs while delaying transition into elongation. *PLoS One*, **9**: e91859.

[14] Ellinger, T. and Ehricht, R. (1998) Singe-step purification of T7 RNA polymerase with a 6-histidine tag. *Biotechniques*., **24**: 718-720.

[15] Chen, T. H., Tanimoto, A., Shkriabai, N., Kvaratskhelia, M., Wysocki, V., and Gopalan, V. (2016) Use of chemical modification and mass spectrometry to identify substrate-contacting sites in proteinaceous RNase P, a tRNA processing enzyme. *Nucleic Acids Res*. **44**: 5344-5355.

[16] Lee, G. H., Lim, H. K., Jung, W., and Hah, S. S. (2012) Incorporation efficiency of 5'-azido-5'-deoxyguanosine into 5'-terminus of RNA for preparation of azido-functionalized RNA. *Bull. Korean Chem. Soc.*, **33**: 3861-3863.

[17] Hofer, K., Abele, F., Schlotthauer, J., and Jaschke, A. (2016) Synthesis of 5'-NAD-capped RNA. *Bioconjug. Chem.*, **27**: 874-877.

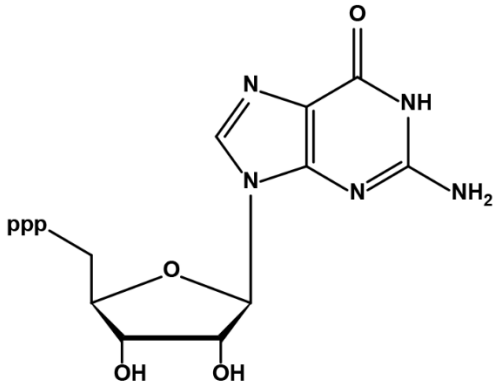
[18] Helmling, C., Keyhani, S., Sochor, F., Furtig, B., Hengesbach, M., and Schwalbe, H. (2015) Rapid NMR screening of RNA secondary structure and binding. *J. Biomol. NMR*, **63**: 67-76.

Chapter 3: A T7 RNA polymerase mutant enhances the yield of 5'-deoxy-5'-azidoguanosine-initiated RNAs

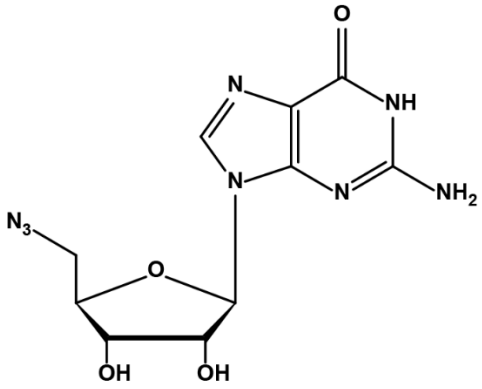
3.1: Background and rationale

While availability to produce 5'-thG-initiated RNA transcripts using T7RNAP P266L will prove useful for many endeavors, other applications such as fluorescence or electron paramagnetic resonance (EPR) studies necessitate the incorporation of chemical handles that in turn will allow covalent tethering to fluors, spin labels, or other probes. The general strategy in these instances entails the use of commercially available probes bearing a reactive group (e.g., aldehyde) that allow for their specific conjugation to a target biomolecule containing an appropriate reactive group (e.g., a primary amine). While there are many approaches to accomplish such site-specific covalent modifications, the azide-alkyne “click reaction” that generates a 1,4-disubstituted five-member 1,2,3-triazole has become a reaction of choice due to its bioorthogonal nature and high efficiency.[1-4] Copper(I) is a catalyst for this 1,3-dipolar cycloaddition, although it does generate reactive oxygen species that degrade RNA. It is notable that RNAs containing a 5'-deoxy-5'-azidoguanosine (^{az}G; **Figure 3.1**) have been successfully ‘clicked’ *in vitro* to Cy5- and fluorescein-alkynes as well as to another RNA having a 3'-alkyne.[2,5,6] The value of ^{az}G for “clicking” to alkynes motivated us to investigate if T7RNAP P266L would enhance the yield of 5'-^{az}G-initiated RNAs akin to the gains that we observed with thG (see **Chapter 2**).

The accurate determination of the percent incorporation of non-fluorescent transcriptional initiators in long RNAs is challenging. Many studies have assessed percent incorporation of a non-fluorescent G-analog based on either the



Guanosine triphosphate



5'-Deoxy-5'-azidoguanosine

Figure 3.1: Structures of guanosine triphosphate and 5'-deoxy-5'-azidoguanosine.

RNA's fluorescence after conjugation of the G-analog to a fluorophore or retardation of the RNA on a polyacrylamide gel after conjugation to biotin/streptavidin.[,6-10] An underlying assumption made in this method is that the reaction between the 5'-G-analog and the fluorophore or biotin proceeds with 100% efficiency, which is often not the case and thus the reported percent incorporation is an underestimate. However, this methodology does not account for non-covalent attachment of fluorophore/biotin to the target RNA, which would overestimate the percent incorporation of the G-analog. An appealing approach to accurately determine percent incorporation of a G-analog at the 5'-terminus of an RNA exploits the difference in electrophoretic mobility of the RNAs bearing either 5'-GTP or 5'-G-analog, made possible by G-analogs that lack the negative charges associated with the triphosphate in the GTP initiator. This method, of course, is most useful for small RNAs where the differences in the 5'-terminus confer an electrophoretic mobility difference that is resolvable in a standard polyacrylamide gel (20 x 20 cm).

To overcome this limitation and accurately determine percent incorporation of ^{az}G in a long RNA, we first employed either T7RNAP WT or P266L in IVTs containing ^{az}G and α-[^{32}P]-UTP to transcribe a pre-tRNA^{Cys} containing a 5-nt, 5'-leader (5'-GGUUU-3'; **Figure 3.2**). Next, we explored using RNase P, an endonuclease that catalyzes the removal of the 5'-leader in pre-tRNAs, to generate two 5 nt leaders: one with ^{az}G at the 5'-terminus and the other with 5'-GTP. These short RNAs (5 nts) with two different termini were easily resolved on a polyacrylamide gel (**Figure 3.3**). Using this approach, we were able to validate our previous finding that T7RNAP P266L decreased abortive transcription that is typically observed when transcription is initiated with G-analogs.

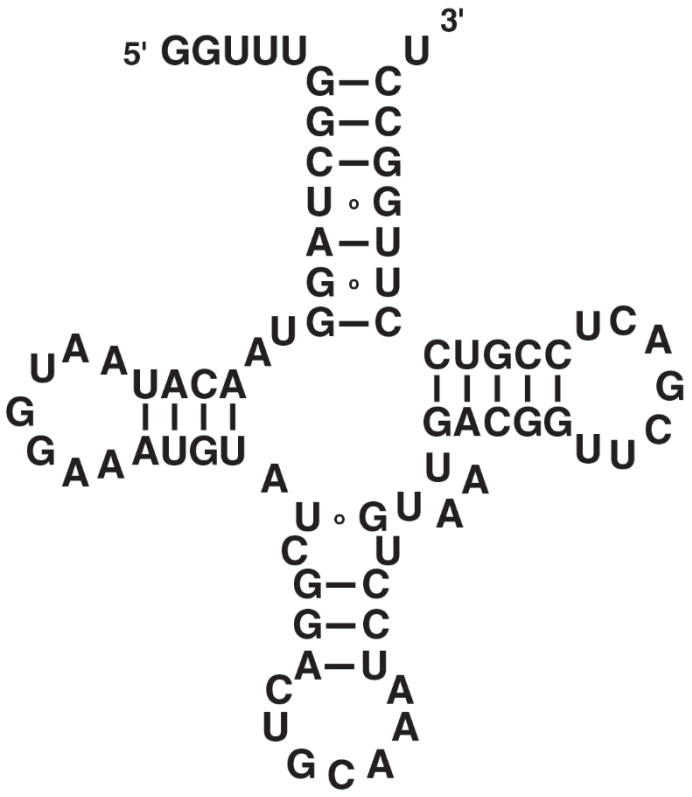


Figure 3.2: Secondary structure of pre-tRNA^{Cys} with a 5-nt 5'-leader.

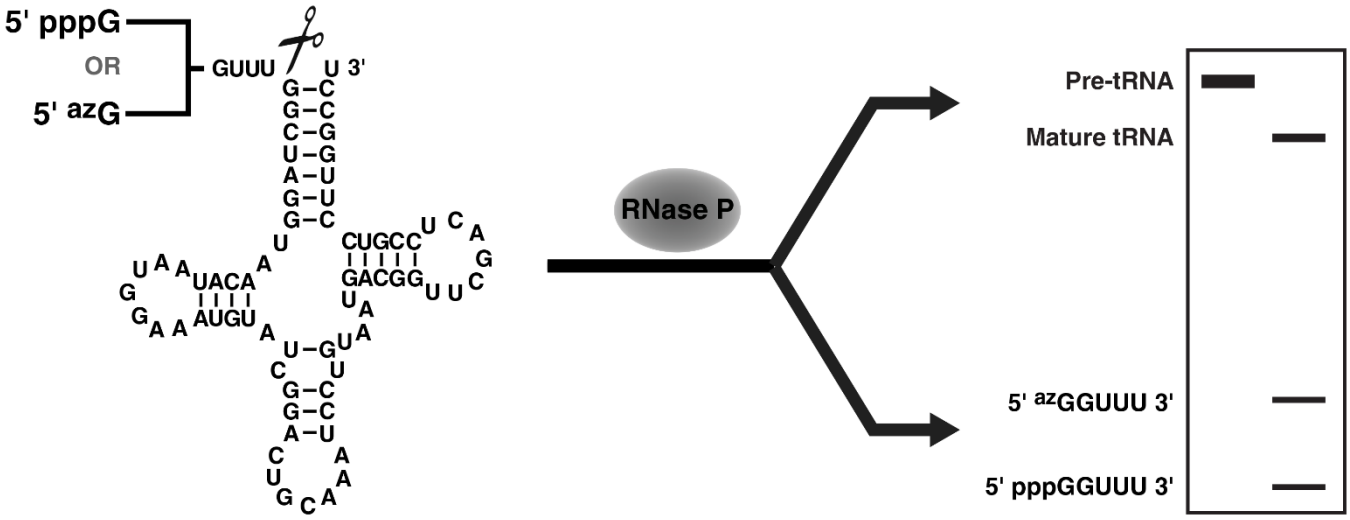


Figure 3.3: Schematic of the RNase P-based assay that was developed to assess the percent incorporation of ^{az}G at the 5'-terminus of pre-tRNA^{Cys}. Upon RNase P cleavage, two different 5-nt leaders are generated: one having 5'-ppp and the other 5'-^{az}G. The 5'-^{az}G leader is expected to migrate slower than the 5'-ppp leader during PAGE owing to its lack of negative charges, thus allowing for direct quantification of percent incorporation of ^{az}G into pre-tRNA^{Cys}.

Importantly, T7RNAP P266L enhances both the total RNA yield and percent incorporation of ^{az}G. Furthermore, we validated that our OPME approach using T7RNAP P266L further enriched for 5'-^{az}G-initiated RNA up to 90%.

3.2: Materials and methods

3.2.1: Overexpression and purification of T7RNAPs WT and P266L

Both WT and P266L T7RNAPs were overexpressed and purified as described earlier (see **Chapter 2.2.1 and 2.2.2**).

3.2.2: In vitro transcription reactions

A PCR DNA template containing a class III Φ6.5 promoter and gene encoding for *A. thaliana* mitochondrial pre-tRNA^{Cys} with a 5-nt 5'-leader and no trailer (**Figure 3.2**) was prepared as previously described.[11] The buffer components, concentrations rNTPs, and amount of T7RNAP WT and P266L in IVTs were the same as those described earlier (see **Chapter 2.2.3 and Figure 2.7**). IVTs containing ^{az}G were performed using 4.8 mM ^{az}G and either 1.2 mM GTP (4:1 ratio) or 0.48 mM GTP (10:1 ratio). One hundred-µL IVTs were performed at 37°C for 5 h. Post-transcription, the DNA template was degraded by addition of 10 U of DNase I (Roche) and the samples were subsequently subjected to phenol-chloroform extraction. Next, the samples were extensively dialyzed at against water at 4°C to remove unincorporated ribonucleotides. The samples were recovered and then precipitated by ethanol precipitation. The

resulting pellets were resuspended in water and total RNA yield was determined using Abs_{260} measurements on a NanoDrop 2000c spectrophotometer and the extinction coefficient of pre-tRNA^{Cys}.

To generate pre-tRNA^{Cys} with an internal radiolabel, 20 µL IVTs were performed as described above, except with the inclusion of 25 µCi of α-[^{32}P]-UTP [3000 mCi/mmol; PerkinElmer, Waltham, MA]. We chose α-[^{32}P]-UTP because the 5-nt 5'-leader of pre-tRNA^{Cys} (5'-GGUUU-3') contains three uridine residues, thus enabling visualization of the 5'-leader after post-PAGE regardless of whether the RNA had been initiated with ^{az}G or GTP. IVTs containing α-[^{32}P]-UTP were performed at 37°C for 3 h and subsequently quenched with an equal volume of loading dye containing 7 M urea, 10 mM EDTA, and 20% (v/v) phenol. The samples were then electrophoresed on an 8% (w/v) polyacrylamide/7 M urea gel at 30 mA for 1 h. Bands corresponding to full-length pre-tRNA^{Cys} were excised and eluted in 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA, 0.01% SDS (w/v). The elution was performed at 37°C for 3 h and then overnight at 4°C. The eluents were filtered through 0.22 µm cellulose acetate Spin-X centrifuge tube filters (Corning Costar, Corning, NY) and then precipitated by ethanol precipitation. The radioactivity of the resulting RNA pellets was counted using a Bioscan QC-4000 XER radioisotope counter (Bioscan, Inc., Washington, D.C.) and the pellets were then resuspended in a small volume of water (typically 10 to 20 µl).

3.2.3: Post-transcriptional removal of 5'-ppp RNAs

The OPME approach was performed essentially as described earlier (see **Chapter 2 section 2.2.3**) but with two exceptions. First, the Xrn-1 incubation was performed at

37°C for 1.5 h. Second, the reactions were performed in a heat block. Total RNA yields of pre-tRNA^{Cys} generated by the OPME approach were not determined because the RNAs were radiolabeled.

3.2.4: RNase P-based assays to determine percent incorporation of ^{az}G

To investigate the relative amounts of pre-RNA^{Cys} initiated with either ^{az}G or GTP, we performed an IVT with α-[^{32}P]-UTP and in the presence of ^{az}G and GTP. The full-length RNA from such an IVT was then cleaved with recombinant *A. thaliana* PRORP1, the protein-only form of RNase P (kindly provided by Dr. Tien-Hao Chen, Gopalan laboratory). The RNase P cleavage assay was performed by incubating trace amounts of pre-tRNA^{Cys} RNA with 2 µM PRORP1 in 20 mM HEPES (pH 7.2), 150 mM NaCl, 1 mM $MgCl_2$, 4 mM DTT, and 5% (v/v) glycerol for 10 min at 24°C.[11] Reactions were quenched by adding an equal volume of loading dye containing 7 M urea, 10 mM EDTA, and 20% (v/v) phenol, and then electrophoresed on a 20% (w/v) polyacrylamide/7 M urea gel (40 cm length x 20 cm width) for 2 h. The gel was then exposed for 3 h to a phosphorimager screen and RNase P cleavage products were visualized using an Amersham Typhoon 5 Biomolecular Imager (GE Healthcare). Percent incorporation of ^{az}G was determined using ImageQuant 5.1 software-based quantitation of the ratio of the slower migrating (RNAs initiated with 5'-^{az}G) to the faster migrating (RNAs initiated with 5'-triphosphate) RNase P product.

3.3: Results

To facilitate comparisons, we used approximately equivalent amounts of T7RNAP WT and P266L in our IVTs (see **Figure 2.7** in **Chapter 2**). While processing internally radiolabeled pre-tRNA^{Cys} that was generated from IVTs using T7RNAP WT, we observed a substantial increase in aborted transcripts when the ^{az}G:GTP ratio was increased from 0:1 to 4:1, and a modest increase when the ratio was increased from 4:1 to 10:1 (**Figure 3.4**). Using an ^{az}G:GTP ratio of 10:1 and either T7RNAP WT or P266L, we found that the ratio of full-length RNA:aborted transcript is significantly higher when T7RNAP P266L is employed in the IVT (**Figure 3.5**). These findings mirror our previous observation that T7RNAP P266L produced fewer abortive transcripts and more full-length RNA than T7RNAP WT when using a 10:1 ratio of thG:GTP (see **Figure 2.8** in **Chapter 2**).

Using a 4:1 ratio of ^{az}G:GTP, we found that T7RNAP WT incorporated ^{az}G into 43% of pre-tRNA^{Cys} transcripts with a total RNA yield of 22 µg from a 100 µL IVT (**Figures 3.6 and 3.9**). Increasing the ^{az}G:GTP ratio to 10:1 and using T7RNAP WT resulted in only a small increase in percent incorporation of ^{az}G into pre-tRNA^{Cys} (52%) while total RNA yield decreased to 10 µg (**Figures 3.6 and Figure 3.9**). In contrast, when T7RNAP P266L was used in an IVT with a 4:1 ratio of ^{az}G:GTP, percent incorporation of ^{az}G into pre-tRNA^{Cys} increased to 60% with a total RNA yield of 42 µg (**Figures 3.7 and 3.9**). After the ^{az}G:GTP ratio was increased 10:1, T7RNAP P266L initiated 75% of pre-tRNA^{Cys} although the yield decreased to 29 µg (**Figures 3.7 and 3.9**). Therefore, T7RNAP P266L yielded ~2.5-fold more 5'-^{az}G-initiated RNA than T7RNAP WT when the ^{az}G:GTP ratio was 4:1 ratio and ~4-fold more 5'-^{az}G-initiated-

RNA than T7RNAP WT when the ^{az}G:GTP ratio was 10:1. Furthermore, using the OPME approach with T7RNAP P266L and a 10:1 ratio of ^{az}G:GTP, we ascertained that ~90% of the RNA containing a 5'-^{az}G (**Figures 3.8 and 3.9**), however, RNA yields were not determined because the RNAs were radiolabeled. That ~10% of the remaining RNA contained a 5'-GTP may reflect sub-optimal functionality of RppH and Xrn-1 were not optimal under our reaction conditions, an inference we had drawn previously from IVTs with thG (see **Chapter 2**).

To ensure that the slower migrating band in the RNase P assays corresponded to 5'-^{az}G-initiated RNA, pre-tRNA^{Cys} that was generated without inclusion of ^{az}G was first cleaved by RNase P to yield the 5'-GTP-bearing 5-nt leader. Calf intestinal phosphatase (CIP) was then used to dephosphorylate the 5'-GTP-bearing 5-nt leader to yield a 5'-OH terminus, which (as expected) migrated similarly as the 5'-^{az}G-containing 5-nt leader (**Figure 3.7**).

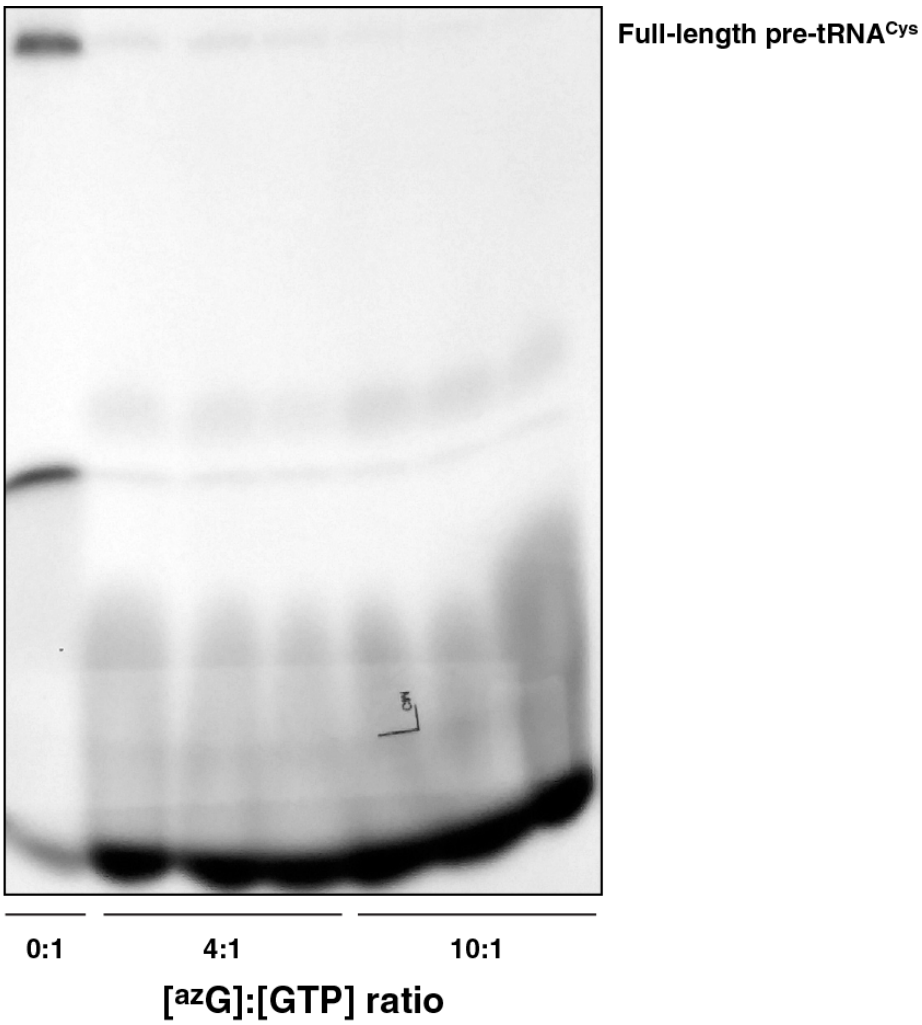


Figure 3.4: Aborted transcription increases with increasing [^{az}G]:[GTP] ratios in IVTs performed with T7RNAP WT. IVTs to generate pre-tRNA^{Cys} were performed using T7RNAP WT with 0:1, 4:1, or 10:1 ratio of ^{az}G:GTP. After transcription was terminated, samples were electrophoresed on an 8% (w/v) polyacrylamide/7 M urea gel at 30 mA for 1 h and full-length pre-tRNA^{Cys} was gel purified. Prior to excising the full-length RNA for gel purification, the gel was exposed to an X-ray film screen and developed in a darkroom. Technical replicates are grouped using an underline below the corresponding lanes.

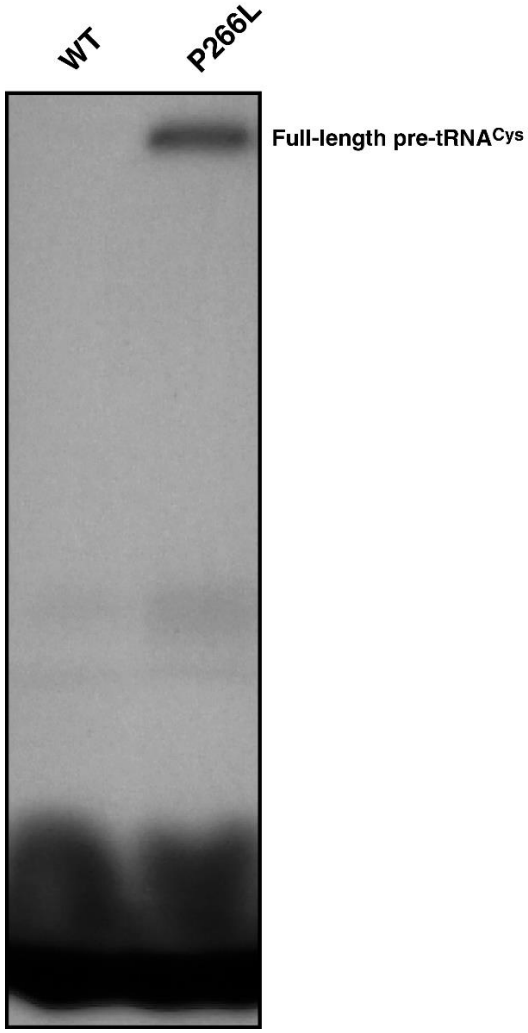


Figure 3.5 T7 RNAP P266L yields fewer aborted transcripts than WT when IVTs are performed with ^{az}G. IVTs to generate pre-tRNA^{Cys} were performed using either T7 RNAP WT or P266L with a 10:1 ratio of ^{az}G:GTP. After transcription was terminated, samples were electrophoresed on an 8% (w/v) polyacrylamide/7 M urea gel at 30 mA for 45 min and the full-length pre-tRNA^{Cys} was gel purified. Prior to excising the full-length RNA for gel purification, the gel was exposed to an X-ray film screen and developed in a darkroom. Brightness and contrast settings were adjusted to ensure minor populations were easily visualized.

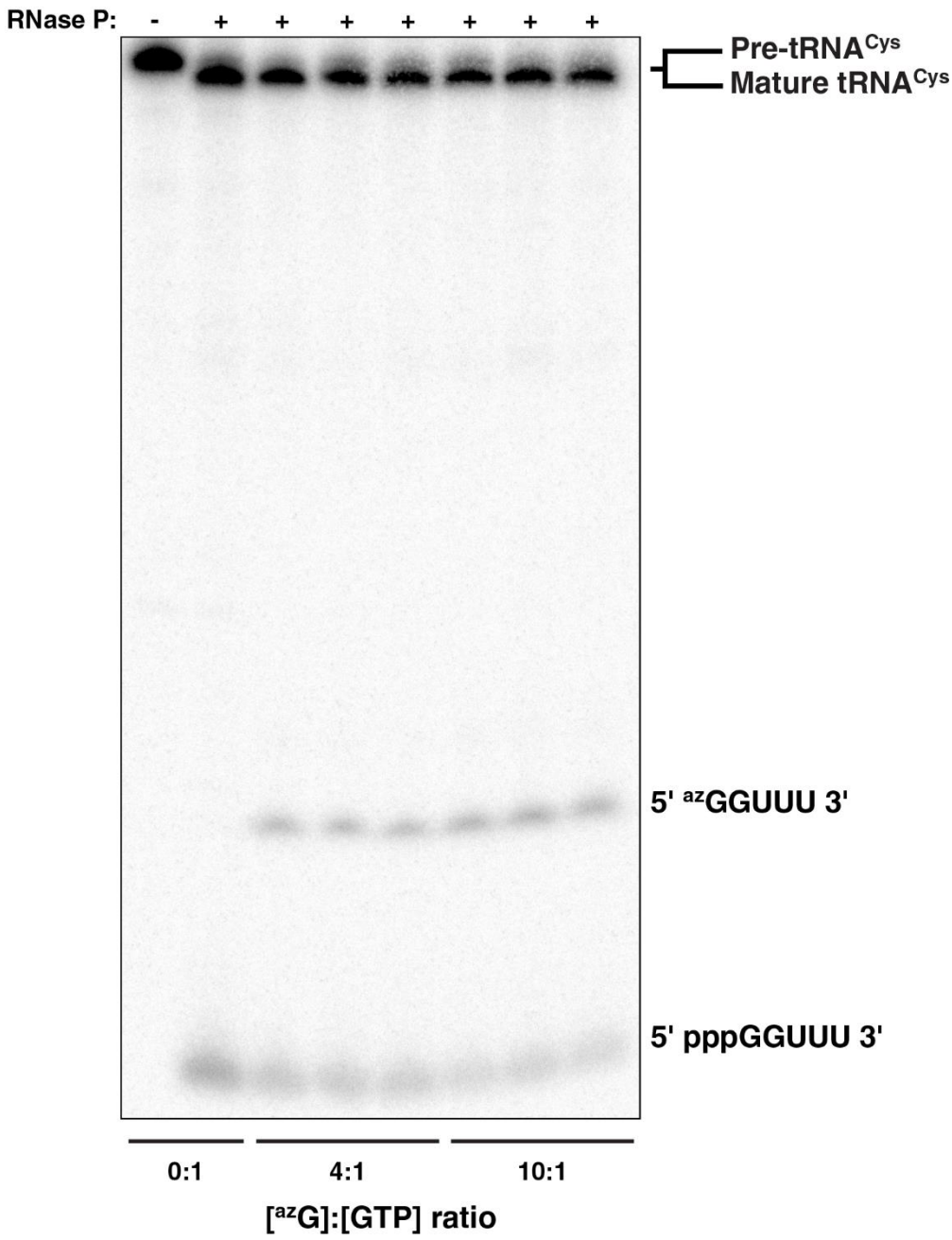


Figure 3.6: RNase P cleavage of pre-tRNA^{Cys} that was generated using T7RNAP WT and a 0:1, 4:1, or 10:1 ratio of ^{az}G:GTP. Technical replicates are grouped using an underline below the corresponding lanes.

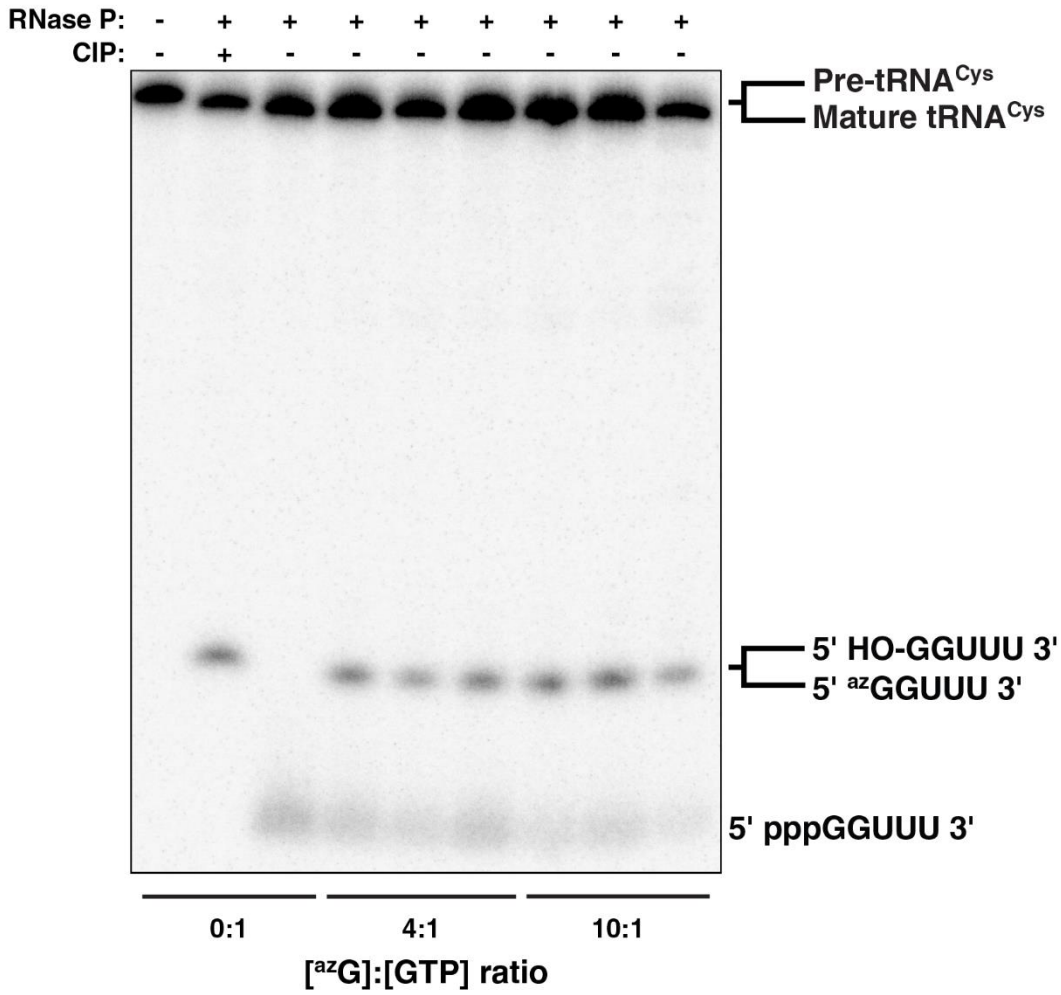


Figure 3.7: RNase P cleavage of pre-tRNA^{Cys} that was generated using T7RNAP P266L and a 0:1, 4:1, or 10:1 ratio of ^{az}G:GTP. Technical replicates are grouped using an underline below the corresponding lanes.

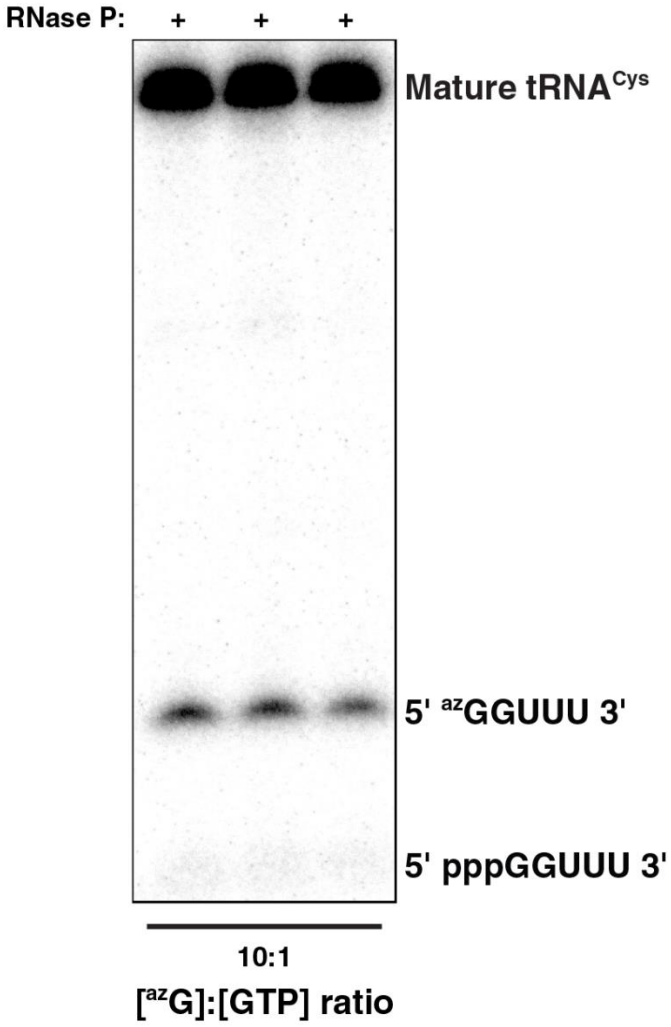


Figure 3.8: RNase P cleavage of pre-tRNA^{Cys} that was generated using the OPME approach with T7RNAP P266L and a 10:1 ratio of ^{az}G:GTP. Technical replicates are grouped using an underline below the corresponding lanes.

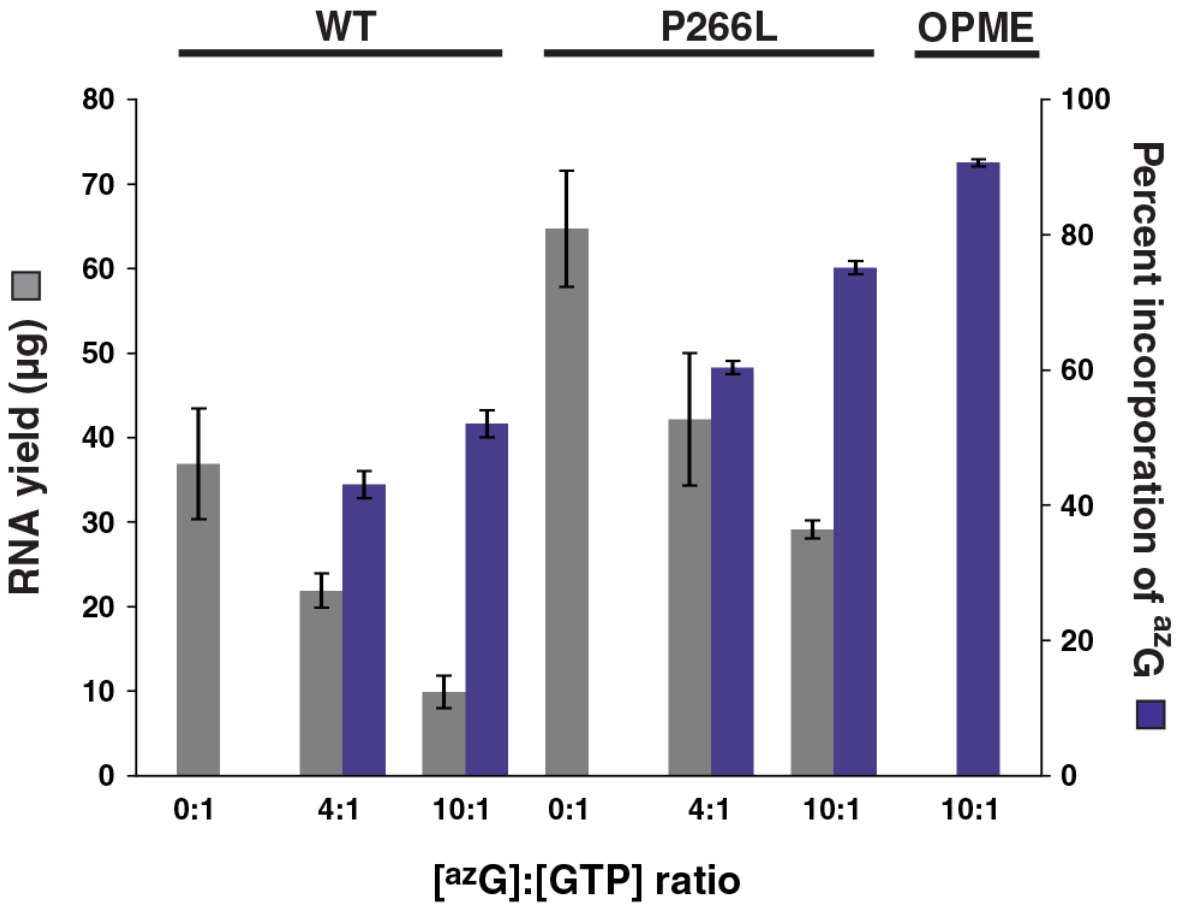


Figure 3.9: RNA yields (gray) and percentage incorporation of ^{az}G (purple) from IVTs containing 0:1, 4:1, or 10:1 ratio of ^{az}G:GTP. IVTS entailed the use of T7RNAP WT or P266L; in the case of the latter, results from the OPME method are also depicted, however, RNA yields were not determined because the RNAs were radiolabeled. Mean and standard deviation values were calculated from three independent measurements.

3.4: Discussion

Our results demonstrate that T7RNAP P266L enhances the total RNA yield and percent of 5'-^{az}G RNA, and that the OPME approach can further enrich the percent of 5'-^{az}G-initiated RNA up to 90%. The findings with ^{az}G parallel those obtained when thG was the G-analog (see **Chapter 2**), underscoring the remarkable tolerance of T7RNAP P266L to generate 5'-G-analog-initiated RNAs. Moreover, the RNase P-based assay offers an accurate method to determine the percent incorporation of 5'-G-analogs in long RNAs. While we used a pre-tRNA as an ideal model substrate and PRORP as the enzyme for the assay, *E. coli* RNase P has been shown to cleave non-tRNA substrates, provided there is an external guide sequence hybridized to the cleavage site (**Figure 3.10**).[12] Thus, any RNA can be assayed for the presence of a modification at the 5'-terminus.

Our results on the incorporation efficiency of ^{az}G into RNA when using T7RNAP WT and a 4:1 ratio of ^{az}G:GTP are in agreement with a previous study, which employed a click reaction between 5'-^{az}G RNA and 6-heptynoyl-*p*-nitroaniline to determine percent incorporation ^{az}G and reported that using a 5:1 ratio of ^{az}G:GTP results in 36% incorporation of ^{az}G into a 75-nt RNA.[6] However, Paredes and Das (2011) concluded that a 4:1 ratio of ^{az}G:GTP resulted in ~100% incorporation of ^{az}G at the 5'-terminus of a 66-nt RNA, as assessed by the small number of unreacted transcripts observed on an ethidium bromide-stained polyacrylamide gel after click ligation of the 5'-^{az}G RNA to another 3'-alkyne RNA.[10]. While we were not successful in achieving such a high incorporation of ^{az}G into any RNA transcribed from either T7RNAP WT or P266L during an IVT, the OPME approach did afford selective enrichment of the modified RNA.

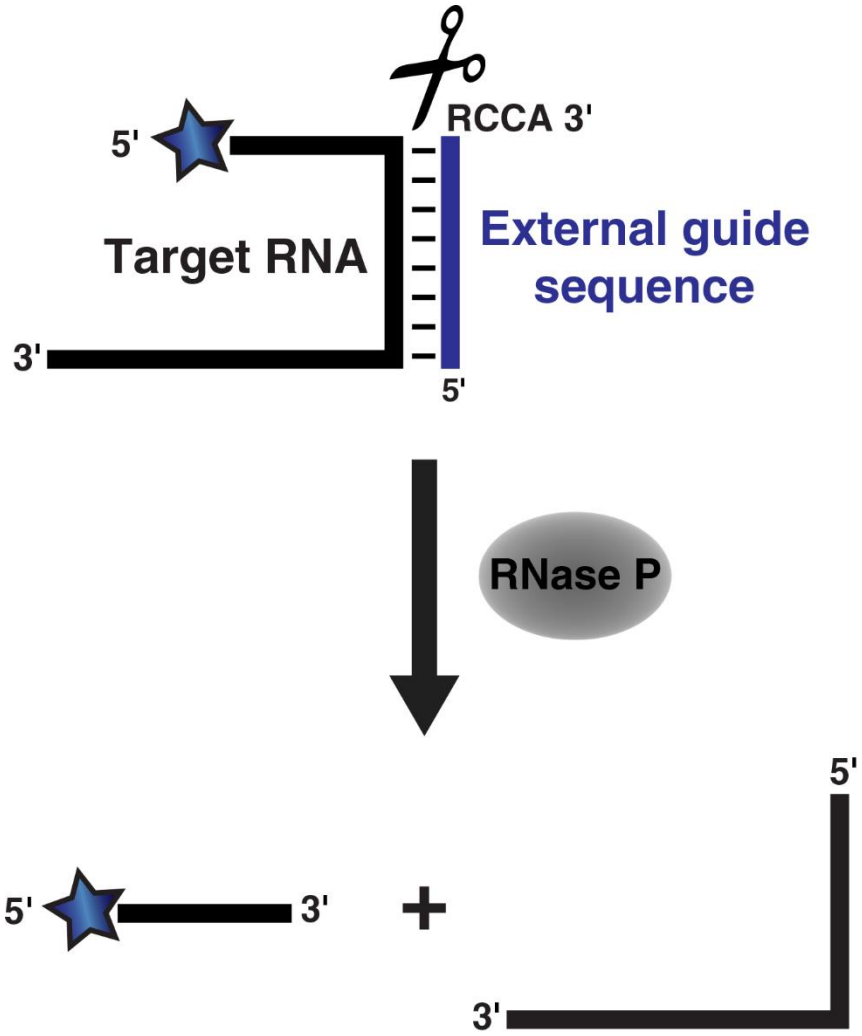


Figure 3.10: Schematic of how any target RNA can be made into a substrate by *E. coli* RNase P after hybridization of an external guide sequence that terminates in a purine (R) and CCA.[12] This approach could be used to determine percent incorporation of a G-analog in any target (see **Figure 3.2**).

3.5: References

[1] Uttamapinant, C., Anupong, T., Grecian, S., Clarke, S., Singh, U., Slade, P., Gee, K. R., and Ting, A. Y. (2013) Fast, cell-compatible click chemistry with copper-chelating azides for biomolecular labeling. *Angew. Chem. Int. Ed.*, **51**: 1199-1216.

[2] Paredes, E., Evans, M., and Das, S. R. (2011) RNA labeling, conjugation and ligation. *Methods*, **54**: 251-259.

[3] Alexander, S. C., and Devaraj, N. K. (2017) Developing a fluorescent toolbox to shed light on the mysteries of RNA. *Biochemistry*, **56**: 5185-5193.

[4] Rombouts, K., Braeckmans, K., and Remaut, K. (2016) Fluorescent labeling of plasmid DNA and mRNA: gains and losses of current labeling strategies. *Bioconj. Chem.*, **27**: 280-297.

[5] Wallace, A. (2013) Fluor-labeling of RNA and fluorescence-based studies of precursor-tRNA cleavage by *Escherichia coli* ribonuclease RNase P. *MS thesis*, The Ohio State University, Columbus, OH.

[6] Lee, G. H., Lim, H. K., Jung, W., and Hah, S. S. (2012) Incorporation efficiency of 5'-azido-5'-deoxyguanosine into 5'-terminus of RNA for preparation of azido-functionalized RNA. *Bull. Korean Chem. Soc.*, **33**: 3861-3863.

[7] Skipsey, M., Hack, G., Hooper, T. A., Shankey, M. C., Conway, L. P., Schroder, M., and Hodgson, D. R. W. (2013) 5'-Deoxy-5'-hydrazinylguanosine as an initiator of T7 RNA polymerase-catalyzed transcriptions for the preparation of labeling-ready RNAs. *Nucleos. Nucleot. Nucl.*, **32**: 670-681.

[8] Zhang, L., Sun, L., Cui, Z., Gottlieb, R. L., and Zhang, B. (2001) 5'-Sulfhydryl-modified RNA: initiator synthesis, *in vitro* transcription, and enzymatic incorporation. *Bioconjug*. *Chem*., **12**: 939-948.

[9] Williamson, D., Cann, M. J., and Hodgson, D. R. W. (2007) Synthesis of 5'-amino-5'-deoxyguanosine-5'-N-phosphoramidate and its enzymatic incorporation at the 5'-termini of RNA molecules. *Chem Commun*., **47**: 5096-5098.

[10] Paredes, E. and Das, S. R. (2011) Click chemistry for rapid labeling and ligation of RNA. *ChemBioChem*, **12**: 125-131.

[11] Chen, T. H., Tanimoto, A., Shkriabai, N., Kvaratskhelia, M., Wysocki, V., and Gopalan, V. (2016) Use of chemical modification and mass spectrometry to identify substrate-contacting sites in proteinaceous RNase P, a tRNA processing enzyme. *Nucleic Acids Res*. **44**: 5344-5355.

[12] Gopalan, V., Vioque, A., and Altman, S. (2002) RNase P: variations and uses. *J. Biol. Chem.*, **277**: 6759-6762.

Chapter 4: Further prospects

4.1: Broader applications of T7 RNA polymerase P266L

Our finding that T7RNAP P266L displays an increased tolerance to initiate RNAs with G-analogs should motivate the study of other applications with this T7RNAP mutant. Many G-analogs are not readily accepted by T7RNAP WT and therefore not widely used despite having favorable chemical properties. Hydrazinylguanosine (^{hyd}G) is one such G-analog that was previously found to be incorporated into only 40% of an 82 nt RNA transcript generated by T7RNAP WT.[1] Incorporation of a hydrazine at the 5'-terminus of an RNA would provide the RNA with a nucleophile that is at least one order more reactive than a primary amine, and therefore, allow for the conjugation of various probes that are available as isothiocyanate or succinimide derivatives.

While T7RNAP can initiate transcripts with either GTP (using a class III Φ6.5 promoter)[2,3] or ATP (using a class II Φ2.5 promoter),[4-6] the polymerase is unable to initiate with a pyrimidine. When RNAs with a pyrimidine at the 5'-terminus are desired, longer transcripts initiated with GTP are generated by IVTs and appropriately placed ribozymes are used to release the downstream, desired RNA with a 5'-pyrimidine.[7] However, this method is time-consuming, not 100% efficient, and lowers total RNA yield given the transcription cost to make the larger RNA. Therefore, another useful application for T7RNAP P266L would be to utilize the enzyme to introduce pyrimidines at the 5'-terminus of an RNA by initiation with a CpG or UpG dinucleotide. Such an IVT would yield a mixture of transcripts: either 5'-GTP- or CpG/UpG-initiated RNA. The OPME approach described in Chapters 2 and 3 could be employed to degrade the 5'-GTP-initiated RNA. Assessing the percent incorporation of the dinucleotide poses a

challenge. However, if a CpthG or UpthG dinucleotide could be synthesized, percent incorporation of the dinucleotide and effectiveness of the OPME approach is easily determined (as shown in **Chapter 2**). Alternatively, post-dephosphorylation of the IVT products, the RNase P-based assay (described in **Chapter 3**) could be employed to resolve 5 and 6 nt RNAs.

4.2: Other potential methods to improve the incorporation of nucleoside analogs

Despite the advances with T7RNAP P266L to initiate RNAs with G-analogs, our total RNA yield is consistently ~two or three-fold lower than the yield from IVTs lacking a triphosphate. Chelliserykattil and Ellington (2004) used a directed-evolution approach to select for T7RNAP mutants that displayed the greatest ability to transcribe fully modified 2'-*O*-methyl RNA and obtained a panel of mutants whose versatility is still under study.[8-10] The most active mutants were identified by their ability to transcribe full-length RNA when one or more rNTPs were replaced with their 2'-*O*-methyl-NTP counterpart.[8] A similar scheme could be applied to select for T7RNAP mutants that enhance the yield of 5'-G-analog initiated RNAs (including ^{az}G and thG). Such tailor-made mutants together with the OPME approach could greatly expand the repertoire of 5'-modifications in RNAs generated by IVTs.

T7RNAP WT has also been reported to successfully accept adenosine analogs as transcriptional initiators when using a class II Φ2.5 promoter,[4-6] although RNA yield from this promoter is 70% lower than that of the class III Φ6.5 promoter.[11] We attempted to use T7RNAP P266L to initiate RNAs with 5'-deoxy-5'-azidoadenosine as

well as isothiazolo[4,3-*d*]-adenosine, however, these attempts were unsuccessful (data not shown). Our results indicate that the advantage conferred by the P266L mutation to enhance the yield of 5'-G-analog initiated RNA when using the class III Φ6.5 promoter does not translate to the ATP-initiating promoter. This finding should prompt efforts to crystallize the initiation complex of T7RNAP bound to the class II Φ2.5 promoter to uncover amino acid residues that are pertinent to the binding of the initiating ATP nucleotide. Subsequent computational studies could suggest potential mutations to T7RNAP that may improve its propensity to initiate RNAs with A-analogs. Alternatively, a directed evolution study could be undertaken to select for T7RNAP mutants that best initiate RNAs with ATP. In this approach, a library of mutant T7RNAPs could be transformed into *E. coli*, whose survival depends on the ability to transcribe a house-keeping RNA (e.g., a tRNA, RNase P) that initiates with an ATP. Advantage-conferring T7RNAP mutations would be selected for based on the growth of the *E. coli* colonies and ensuing IVT tests with an A-analog could be used to identify mutants that enhance the yield of 5'-A-analog initiated RNA.

4.3: References

[1] Skipsey, M., Hack, G., Hooper, T. A., Shankey, M. C., Conway, L. P., Schroder, M., and Hodgson, D. R. W. (2013) 5'-Deoxy-5'-hydrazinylguanosine as an initiator of T7 RNA polymerase-catalyzed transcriptions for the preparation of labeling-ready RNAs. *Nucleos. Nucleot. Nucl.*, **32**: 670-681.

[2] Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.*, **180**: 51-62.

[3] Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res*., **15**: 8783-8798.

[4] Huang, F., He, J., Zhang, Y., and Guo, Y. (2008) Synthesis of biotin-AMP conjugate for 5' biotin labeling of RNA through on-step *in vitro* transcription. *Nat. Protoc.*, **3**: 1848-1861.

[5] Huang, F., and Shi, Y. (2010) Synthesis of symmetrical thiol-adenosine conjugate and 5' thiol-RNA preparation by efficient one-step transcription. *Bioorg. Med. Chem. Lett.*, **20**: 6254-6257.

[6] Li, N., Yu, C., and Huang, F. (2005) Novel cyanine-AMP conjugates for efficient 5' RNA fluorescent labeling by one-step transcription and replacement of [γ-^{32}P]ATP in RNA structural investigation. *Nucleic Acids Res*., **33**: e37.

[7] Ahmed, Y. L., and Ficner, R. (2014) RNA synthesis and purification for structural studies. *RNA Biol.*, **11**: 427-432.

[8] Chelliserrykattil, J., and Ellington, A. D. (2004) Evolution of a T7 RNA polymerase variant that transcribes 2'-O-methyl RNA. *Nat. Biotechnol.*, **22**: 1155–1160.

[9] Meyer, A. J., Garry, D. J., Hall, B., Byrom, M. M., McDonald, H. G., Yang, X., Yin, Y. W., and Ellington, A. D. (2015) Transcription yield of fully 2'-modified RNA can be increased by the addition of thermostabilizing mutations to T7 RNA polymerase mutants. *Nucleic Acids Res.*, **43**: 7480–7488.

[10] Kimoto, M., Meyer, A. J., Hirao, I., and Ellington, A. D. (2017) Genetic alphabet expansion transcription generating functional RNA molecules containing a five-letter alphabet including modified unnatural and natural base nucleotides by thermostable T7 RNA polymerase variants. *Chem. Commun.*, **53**: 12309–12312.

[11] Sun, G., and Riggs, A. D. (2017) A simple and cost-effective approach for *in vitro* production of sliced siRNAs as potent triggers for RNAi. *Mol.Ther. Nucleic Acids*, **8**: 345–355.